

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 12:58:43 ; Search time 85 Seconds
(without alignments)

1641.714 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 1938

Sequence: 1 MRSALLSAVALLRAQQPF.....LRHNRRLRPLPTLFRNLI 389

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : A_Geneseq_23Sep04;*

1: geneseqp1980s;*
2: geneseqp1990s;*
3: geneseqp2000s;*
4: geneseqp2001s;*
5: geneseqp2002s;*
6: geneseqp2003as;*
7: geneseqp2003bs;*
8: geneseqp2004s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1938	100.0	566	3	AYAY69183		Aay69183 Amino aci
2	1938	100.0	566	5	AAU78891		Aau78891 Mouse end
3	1477	76.2	560	2	AAR71294		Aar71294 Human gly
4	1477	76.2	560	3	AYAY69184		Aay69184 Amino aci
5	1477	76.2	560	5	AAU78892		Aau78892 Human end
6	1477	76.2	560	7	ADF69095		Adf69095 Human MP5
7	607	31.3	581	5	ABB53264		Abb53264 Human pol
8	607	31.3	581	6	ABU52615		Abu52615 Human NOV
9	607	31.3	581	6	ABU56706		Abu56706 Lung canc
10	607	31.3	581	6	ABU56544		Abu56544 Lung canc
11	607	31.3	581	6	ABR43175		Abr43175 Human REM
12	607	31.3	581	7	ADC83709		Adc83709 Human leu
13	607	31.3	581	7	ADE08191		Ade08191 Novel pro
14	607	31.3	581	7	ADN39956		Adn39956 Cancer/an
15	607	31.3	581	7	ADN39092		Adn39092 Cancer/an
16	607	31.3	790	7	ADC86793		Adc86793 Human GPC
17	607	31.3	819	6	ABP97224		Abp97224 Tumour-as
18	583	30.1	578	7	ADC83708		Adc83708 Rat leuci
19	547	28.2	444	7	ADE07921		Ade07921 Novel pro
20	509.5	26.3	395	6	ABU12114		Abu12114 Human pro
21	503.5	26.0	570	8	ADI36915		Adi36915 Human LRR
22	445	23.0	545	5	ABP64705		Abp64705 Human pro
23	445	23.0	545	6	ABJ26653		Abj26653 Human pro
24	445	23.0	545	6	ABU54560		Abu54560 Human NOV
25	445	23.0	545	6	ABU12117		Abu12117 Human pro

The present sequence represents a murine glycoprotein V (GP V). The specification describes non-human transgenic animals which contain or comprise a modified GP V gene. GP V is a platelet and endothelial cell specific glycoprotein, and is a substrate for thrombin. Activation of platelets by thrombin results in the loss of surface GP V. Platelets from the transgenic animals can be used in a number of assays to identify agents that modulate GP V function, or to assess the role of GP V in platelet function. Such assays may help elucidate the extent to which GP

RESULT 1
ID AAY69183 Standard; protein; 566 AA.
XX
AC AAY69183;
DT 30-MAY-2000 (first entry)
XX
DE Amino acid sequence of murine glycoprotein V (GP V).
XX
KW Glycoprotein V; GP V; transgenic animal; platelet; thrombin substrate; platelet function; haemostasis; platelet aggregation; anticoagulant; thrombolytic; antiplatelet therapeutic drug.
XX
OS Mus sp.
XX
FH Key Misc-difference 566
FT Location/Qualifiers
FT /note= "encoded by TG"
XX
PN WO200008137-A2.
XX
PD 17-FEB-2000.
XX
PP 04-AUG-1999; 99WO-US017594.
XX
PR 04-AUG-1998; 98US-010979P.
XX
(CORT-) COR THERAPEUTICS INC.
XX
PA Ramakrishnan V, Phillips DR;
XX
PI WPI; 2000-195570/17.
XX
DR N-PSDB; AAZ61216.
XX
PT Transgenic animals with a modified glycoprotein V gene used to identify agents that modulate the effects of glycoprotein V.
XX
PS Example 1; Fig 2; 44pp; English.

The present sequence represents a murine glycoprotein V (GP V). The specification describes non-human transgenic animals which contain or comprise a modified GP V gene. GP V is a platelet and endothelial cell specific glycoprotein, and is a substrate for thrombin. Activation of platelets by thrombin results in the loss of surface GP V. Platelets from the transgenic animals can be used in a number of assays to identify agents that modulate GP V function, or to assess the role of GP V in platelet function. Such assays may help elucidate the extent to which GP

CC V is critical for normal haemostasis. The transgenic animals may also be
 CC used to identify agents that modulate platelet aggregation. Agents that
 CC can be tested include anticoagulants, thrombolytics, and antiplatelet
 CC therapeutic drugs
 XX Sequence 566 AA;

Query Match 100.0%; Score 1938; DB 3; Length 566;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Db 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Qy 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Db 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Qy 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Db 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Qy 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Db 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Qy 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Db 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Qy 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Db 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Qy 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389
 Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 2
 ID AAU78891 standard; protein; 566 AA.
 XX AAU78891;
 XX DT 02-JUL-2002 (first entry)
 DE Mouse endothelial cell specific glycoprotein GP V.

XX Endothelial cell specific glycoprotein; GP V; anticoagulant;
 KW platelet activation inhibitor; thrombin-induced activity;
 KW transgenic animal; anti-thrombotic agent; arterial thrombosis;
 KW atherosclerotic artery; acute myocardial infarction; stroke;
 KW ischaemic complication of cardiovascular disease;
 KW Bernard-Soulier syndrome; BSS; mouse.
 XX Mus sp.
 XX WO200217711-A2.
 XX PD 07-MAR-2002.
 XX PF 31-AUG-2001; 2001WO-US026936.
 XX PR 31-AUG-2000; 2000US-0229047P.
 PR 31-AUG-2000; 2000US-0230566P.
 XX PA (CORT-) COR THERAPEUTICS INC.
 XX PI Ramakrishnan V, Phillips D;
 XX DR 2002-315494/35.

DR N-PSDB; ABK48274.
 XX Identifying agent inhibiting thrombin-induced platelet activation, by
 PT administering test agent and proteolytically inactive thrombin to GP V
 PT null non-human transgenic animal and monitoring platelet aggregation.
 XX
 SQ Example 1; Fig 2; 64PP; English.
 XX The invention describes a method of identifying an agent that inhibits
 CC thrombin-induced activity (e.g., platelet activation). Thrombin-induced
 CC activity can be modulated (inhibited) by endothelial cell specific
 CC glycoprotein GP V, by administering test agent and (proteolytically
 CC inactive) thrombin to a GP V null non-human transgenic animal (I), and
 CC monitoring aggregation of platelets in (I) to identify inhibition. The
 CC method is useful for identifying agents that inhibit thrombin-induced
 CC activity e.g. thrombin-induced platelet activation, and for screening
 CC anti-thrombotic agents. Inhibitors of thrombin-induced activity are
 CC useful for inhibiting arterial thrombosis in atherosclerotic arteries
 CC e.g. acute myocardial infarction and stroke, inducing ischaemic
 CC complications of cardiovascular disease and platelet activation in a
 CC subject and may be useful for treatment of Bernard-Soulier syndrome
 CC (BSS). This is the amino acid sequence of the murine endothelial cell
 CC specific glycoprotein GP V isolated in the invention
 XX Sequence 566 AA;
 Query Match 100.0%; Score 1938; DB 5;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Db 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Qy 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Db 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Qy 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Db 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Qy 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Db 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Qy 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Db 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Qy 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Db 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Qy 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389
 Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

Query Match 100.0%; Score 1938; DB 5;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Db 1 MIRSALLSAVALLRAQPFPCPKTCKCVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 Qy 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Db 61 DQGILRNHSFSGMTVLQRIMLSDSHISAIDPGTFNDLVKLTKLRLTRNKISRLPRAILDK 120
 Qy 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Db 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNLSELFPANLFFSLRELKLLDLSRN 180
 Qy 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Db 181 NLTHLPKGLLGAAQVKLEKLLYNSNQLTSVDSGLLSNLGALTELRLERNHLSRSVAPGAFDR 240
 Qy 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Db 241 LGNLSSLTLSGNLLESLPPALFLHVSSVRLTLPENPPEELPDVLFGE MAGLREWLNGT 300
 Qy 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Db 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFCQGLRELRVLAHTNALAELRDDALR 360
 Qy 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389
 Db 361 GLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

RESULT 3
 ID AAR71294
 XX AAR71294 standard; protein; 560 AA.
 AC AAR71294;
 XX
 DE Human glycoprotein V.
 KW Glycoprotein V; GPV; platelet.
 KW Homo sapiens.

XX	SQ	Sequence 560 AA;	Query Match	Score 1477;	DB 3;	Length 560;
			Best Local Similarity 76.1%;	Pred. No. 1e-127;		
			Matches 296;	Conservative 37;	Mismatches 56;	Indels 0;
					Gaps 0;	
QY	1	MLRSSALLSAVALLRAQPFPCPKTCVKVVRDAACQCSGGSVAHIAELGLPTNLTHILLFRM 60				
	1	MLRGTLCAVLLCAVLLCAVLLRAQPFPCPPACKCVRDAACQCSGGDVARIISALGLPTNLTHILLFGM 60				
	61	DQGILRNHHSFSGMTVLRQLMISDISHISAIIDPGTFNDLVKLKTKLRLTRNKKISRLPRAILDK 120				
	61	GRGVHQSQSFSGMTVLRQLMISDISHISAVAPGTFSDLIKLKTKLRLSRNKKITHLPGALLDK 120				
	121	MVLEQQLFLDHNALRDLQDNLFOQLRNLQELGLNQNQLSFLPANLFSSLRELKLLDLSRN 180				
	121	MVLEQQLFLDHNALRQIDQNMFQKLVNLQELALNQNQLDLFLPASLFTNLLENLKLLDLSGN 180				
	181	NLTHLPKGGLLGAQVKLEKLLLYSNQNLTSVDSGLSNLGAALTLLRERNHLSVAPGAFDR 240				
	181	NLTHLPKGGLLGAQAKLERLLHSNRLVSLDSGLLNSLGAALTLLQFHRNHIRSIAAPGAFDR 240				
	241	LGNLSSLTLSGNLLESPPALFLHVSSVRLLTFENPYLEELPDPVLFGE MAGLRELWNGT 300				
	241	LPNLISSLTLSRNHNLAFLPSALFLHSHNLTLTFENPLAELPGVLFGE MGGLQELWLNR 300				
	301	HLSTLPAAAFRNLSGLQTGLITRNPRLSALPRGVFQGLRELRLVLAHLTNALAEI RDALAR 360				
	301	QLRTLPAAFRNLSRLRYLGVTLSPLRSALPQGAFQGLGELQVLALHSNGLTALPDGLLR 360				
	361	GLGHLRQVSLRHNRRLRALPRTLFRNLSL 389				
	361	GLGHLRQVSLRRNRLRALPRAFLRNLSL 389				

XX WO200217711-A2.
XX PN
XX 07-MAR-2002.
XX PFD
XX 31-AUG-2001; 2001WO-US026936.
XX PF
XX 31-AUG-2000; 2000US-0229047P.
XX PR
XX 31-AUG-2000; 2000US-0230566P.

RESULT 6
ADDF69095
IDD ADF69095 standard; protein; 560 AA.
XXX
AC
XXX
DT
12-FEB-2004 (First entry)

CC of testes including testicular cancer, male reproductive diseases
 CC including low testosterone and male infertility; and disease of pancreas
 CC including diabetic ketoacidosis, Type 1 and 2 diabetes and obesity. The
 CC present sequence is a polypeptide of the invention
 XX Sequence 581 AA;

Query Match 31.3%; Score 607; DB 5; Length 581;
 Best Local Similarity 37.7%; Pred. No. 4.4e-47;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
 21 CPKTKCVRDAAQCSEGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVQLRM 80
 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLOQILNTHITELNESPPFLNISALIALR 83
 81 LSDSHISAIDPGTENDLYVKLTKLTLTRNKISRLPRAILDRLMVLLEQLFLDENALRDLQDN 140
 84 IERNELSRITPGAFRNGLQSLRYLSLANKLQVLPIGLFQGLDSLESLLSSNQLLQIOPA 143
 141 LFOQLRNLOELGLNQNONSFLPANLFSSLLREKLKLUDLSRNRLNLTLPKGLLGAQVKLEKLL 200
 144 HFSQCSNLKELQHGNHLEYIPDGAFDHLVGLTKLNQKNSLTHISPRVYQHQLGNLQVLR 203
 201 LYSNQLTSDVDSGLLSNLGALETLERLERNHLRSVAPGAFDRLGNLSSLLTSGNLLSPLPA 260
 204 LYENRLTDIPMGTEFDGLVNLQQLNQQLQGLNQGLLSPGLFHNHHNQRLYLISQLPPS 263
 261 LFLHVVSSVSRLTLPENPYLEELPDVLFGEAGLRELMNLGTHLSTLPAAFFRNLSGLQTLG 320
 264 IFMQLPQLNRLTIFGNSLKELSLIGIFGPMPLNLRELLWLYDNHISLIPDNVFSNLRQLVLI 323
 321 LTRNPLRSALPGRVFOGLLRELRLVLAHTNALAELRDDALRGLGLRQVSLSLHNRRLALPR 380
 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDGNGVFRMLANLQNLISLQNNRLRQPG 382
 381 TLFRNLSSL 389
 383 NIFANVNGL 391

RESULT 8
 ABU52615 standard; protein; 581 AA.
 ID ABU52615
 XX AC ABU52615;
 XX DT 04-MAR-2003 (first entry)
 XX DE Human NOVX protein, NOV21.
 XX KW Human; immunogen; NOVX; metabolic disorder; diabetes; cardiomyopathy;
 KW obesity; infectious disease; anorexia; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; immune disorder;
 KW haematopoietic disorder; dyslipidaemia; metabolic disturbance;
 KW metabolic syndrome X; wasting disorder; cancer; gene therapy; SNP;
 KW single nucleotide polymorphism.
 XX OS Homo sapiens.
 XX PH Key Location/Qualifiers
 FT Misc-difference 11 /note= "May be Ala as a result of a single nucleotide polymorphism"
 FT FT Misc-difference 148 /note= "May be Arg as a result of a single nucleotide polymorphism"
 FT FT Misc-difference 430 /note= "May be Gly as a result of a single nucleotide polymorphism"
 FT FT Misc-difference 495 /note= "May be Asn as a result of a single nucleotide polymorphism"
 FT FT XX SQ Sequence 581 AA;

PR WO200281518-A2.
 XX 17-OCT-2002.
 XX 21-FEB-2002; 2002WO-US005374.
 XX PR 21-FEB-2001; 2001US-0270220P.
 PR 21-FEB-2001; 2001US-0270523P.
 PR 23-FEB-2001; 2001US-0270797P.
 PR 08-MAR-2001; 2001US-0274295P.
 PR 16-MAR-2001; 2001US-0276400P.
 PR 16-MAR-2001; 2001US-0276677P.
 PR 26-MAR-2001; 2001US-0278796P.
 PR 04-APR-2001; 2001US-0281521P.
 PR 25-APR-2001; 2001US-0286548P.
 PR 17-MAY-2001; 2001US-0291765P.
 PR 10-AUG-2001; 2001US-0311595P.
 PR 13-AUG-2001; 2001US-0311980P.
 PR 10-SEP-2001; 2001US-0318526P.
 PR 17-SEP-2001; 2001US-0322712P.
 PR 18-OCT-2001; 2001US-03330307P.
 XX (CURA-) CURAGEN CORP.
 XX PI Pena CEA, Shimkets RA, Li L, Shenoy SG, Kekuda R, Spytek KA;
 PI Vernet CAM, Malyankar UL, Guo X, Gusev VY, Casman SJ, Boldog FL;
 PI Furtak K, Tchernev VT, Paturajan M, Gangolli EA, Padigaru M, Liu X;
 PI Baumgartner JC, Gerlach VL, Spaderna SK, Zerhusern BD;
 XX DR WPI; 2003-046859/04.
 XX DR N-PSDB; ABX70670.
 XX PT New isolated NOVX polypeptide useful for treating cardiomyopathy,
 atherosclerosis, metabolic disorders, diabetes, obesity, infectious
 disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
 cancer.
 XX PS Claim 1; Page 179; 479pp; English.
 XX CC The invention relates to an isolated polypeptide termed NOV1, 2a,
 CC 2b, 3a, 3b, 4a, 4b, 5, 6, 7a-c, 8a-e, 9a-b, 10, 11, 12a-c, 13, 14,
 CC 15, 16a-d, 17a-b, 18, 19, 20a-b, 21-30) appearing as ABU52578-ABU52624), a
 CC variant of NOVX, a mature form of NOVX, and a variant of the mature form
 CC of NOVX. Also included are a nucleic acid molecule (NOVX NA) encoding
 CC NOVX, or a fragment or complement of NOVX NA, a vector comprising NOVX
 CC NA, a cell comprising the vector, an anti-NOVX antibody (ab), determining
 CC the presence or amount of NOVX or NOVX NA in a sample, and identifying an
 CC agent that binds or modulates the expression or activity of NOVX. NOVX,
 CC NOVX NA or ab is useful for treating or preventing a NOVX-associated
 CC disorder in a subject, preferably human. Ab is useful for identifying the
 CC presence or amount of NOVX in a sample. NOVX is useful for identifying an
 CC agent that binds to NOVX. NOVX, NOVX NA or ab is useful for treating
 CC metabolic disorders, diabetes, cardiomyopathy, obesity, infectious
 CC disease, anorexia, neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, immune disorders, haematopoietic disorders, and
 CC various dyslipidaemias, metabolic disturbances associated with obesity,
 CC the metabolic syndrome X and wasting disorders associated with chronic
 CC diseases, various cancers, endocrine, connective tissue, blood, vascular,
 CC skin, renal, bone, brain, muscle disorders, or bacterial, fungal,
 CC protozoal or viral infections. NOVX, NOVX NA or ab is useful in screening
 CC assays, detection assays, predictive medicine, and in methods of
 CC treatment. NOVX is useful as immunogen, to screen for potential
 CC ant/agonist compounds, and as bait protein in a two-hybrid or three-
 CC hybrid assay. NOVX NA is useful in gene therapy, to express NOVX, to
 CC detect NOVX mRNA or a genetic lesion in a NOVX gene, and to modulate NOVX
 CC activity. The cell is useful for producing non-human transgenic animals.
 CC Ab is useful for isolating, and purifying NOVX and to monitor protein
 CC levels in tissue as part of a clinical testing procedure. The present
 CC sequence represents a NOVX protein
 XX SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
Db 21 CPKTKCVYRDAACGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLRQLM 80
25 CPSECTCSRASQVECTGARIYAV-PTPLPWNAMSILQILNTHTELNESPFLNISALIALR 83
Qy 81 LSDSHISAIIDPGTFNDLVKLTKLTRNKSIRLPRAILDKMVILLEQLFDHNAIRDDQN 140
Db 84 IEKNELSRTPGAFRNLGSRLYRSLANNKQLVLPQGLFQGLDSLESLLSSNQLLQIOPA 143
Qy 141 LFQQLRNQELGLNQNQLSFLPANLFSLREKLIDLSRNNLTHLPKGGLIGAQVKLEKLL 200
Db 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKLNLGKNSLTHISPRVFOHLGNLQVLR 203
Qy 201 LYSNQLTSVDSGGLLSNLGALTELERNHLSVAFGAFDRLGNLSSLTLSGNLLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSFGLFHANNHNLQRLYLSNNHISQLPPS 263
Qy 261 LFLHVSYSVSRLTTFENPLEELPDVLFGEAGLRELLWLNQGTHLSTLTPAAAFRNLSGLQTLG 320
Db 264 IFMQLPQLNRLLTFLFGNSLKELSLQIGFPMPNLRQDNHISSLPDLVFSNLRQLQVLI 323
Qy 321 LTRNPRLSALPRGVFOGLRELRLTNAELRDLGHLRQVSIRHNRRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDGNVFRMLANLQNISLQNLRLRQLPG 382
Qy 381 TLFRNLSL 389
Db 383 NIFANVNGL 391

RESULT 9
ID ABUS6706 Standard; protein; 581 AA.
XX AC ABUS6706;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #299.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
DR N-PSDB; ABX76435.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX PS Claim 27; Page 423; 453pp; English.

CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences of the
CC invention
XX SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
Db 21 CPKTKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLRQLM 80
25 CPSECTCSRASQVECTGARIYAV-PTPLPWNAMSILQILNTHTELNESPFLNISALIALR 83
Qy 81 LSDSHISAIIDPGTFNDLVKLTKLTRNKSIRLPRAILDKMVILLEQLFDHNAIRDDQN 140
Db 84 IEKNELSRTPGAFRNLGSRLYRSLANNKQLVLPQGLFQGLDSLESLLSSNQLLQIOPA 143
Qy 141 LFQQLRNQELGLNQNQLSFLPANLFSLREKLIDLSRNNLTHLPKGGLIGAQVKLEKLL 200
Db 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLTKLNLGKNSLTHISPRVFOHLGNLQVLR 203
Qy 201 LYSNQLTSVDSGGLLSNLGALTELERNHLSVAFGAFDRLGNLSSLTLSGNLLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSFGLFHANNHNLQRLYLSNNHISQLPPS 263
Qy 261 LFLHSSVSRLTTFENPLEELPDVLFGEAGLRELLWLNQGTHLSTLTPAAAFRNLSGLQTLG 320
Db 264 IFMQLPQLNRLLTFLFGNSLKELSLQIGFPMPNLRQDNHISSLPDLVFSNLRQLPG 323
Qy 321 LTRNPRLSALPRGVFOGLRELRLTNAELRDLGHLRQVSIRHNRRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDGNVFRMLANLQNISLQNLRLRQLPG 382
Qy 381 TLFRNLSL 389
Db 383 NIFANVNGL 391

RESULT 10
ID ABUS6544
XX AC ABUS6544;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #137.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW bronchiectasis; bronchitis; chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
DR N-PSDB; ABX76435.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the

KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;	QY	261 LFLHVSSVSRLLTFENPYLEELPDVLFGE MAGLRELWLN GTHLSTLPAAAFRNL SGLQTLG 320
KW	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.	Db	264 IFLNQLPQLNRLTLEFGNSLKEBSLGLGIPGPMPNRLREBLWLYDNHISSLPDDNFSNLRQLQVLI 323
OS	Unidentified.	QY	321 LTRNPLRLSALP RGVFOGLRRLVLAELRDDALRGLGHLRQVSLLRHNRLRALPR 380
XX	XX	Db	324 LSRN-QISFISPGAFNGLTELRELLSHTNALQDLDGNVFRMLQNLQNLISLQNNRQLRQLPG 382
PN	WO200286443-A2.	QY	381 TLFRNLSSL 389
XX	31-OCT-2002.	Db	383 NTFANTNGL 391
PD	PF 18-APR-2002; 2002WO-US012476.	RESULT 11	
XX	XX	ABR43175	
PR	18-APR-2001; 2001US-0284770P.	ID	ABR43175 standard; protein; 581 AA.
PR	10-MAY-2001; 2001US-0290492P.	XX	
PR	09-NOV-2001; 2001US-0339245P.	AC	ABR43175;
PR	13-NOV-2001; 2001US-0350666P.	XX	
PR	29-NOV-2001; 2001US-0334370P.	DT	07-JUL-2003 (first entry)
PR	12-APR-2002; 2002US-0372246P.	XX	
XX	(EOSB-) EOS BIOTECHNOLOGY INC.	DE	Human REMAP-7 protein SEQ ID NO:7.
PA	XX	XX	
PI	Aziz N, Murray R;	XX	
XX	WPI: 2003-093161/08.	XX	
DR	DR N-PSDB; ABX76272.	XX	
XX	Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.	PT	Human; receptor and membrane-associated protein; REMAP; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV; antiinflammatory; anti-allergic; thyromimetic; gene therapy; atherosclerosis; neurological disorder; epilepsy; Huntington's disease; stroke; AIDS; immune disorder; inflammatory disorder; allergy; developmental disorder; hypothyroidism; Cushing's syndrome; infection.
PS	Claim 27; Page 294; 453pp; English.	XX	
XX	The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a sequence that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. arelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences of the ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention.	CC	
XX	Sequence 581 AA;	SQ	
XX	Score 607; DB 6; Length 581;	PT	Warren BA, Gietzen KJ, Lal PG, Xu Y, Tran UK, Lee S; Borowsky ML, Becha SD, Sanjanwala MM, Tang YT; Honchell CD, Yue H, Elliott VS, Richardson TW, Azimzai Y, Chawla NK; Baughn MR, Lu DAM, Nguyen DB, Kalaus DP, Sapperstein SK; Ramkumar J, Lehr-Mason PM, Griffin JA, Duggan BM, Lee SY; Zebarjadian Y, Hafalia AJA, Gururajan R, Swarnakar A, Lee EA; Marquis JP, Khare R, Emerling BM, Jiang X, Jackson AA; WPI; 2003-354596/33.
XX	Best Local Similarity 37.7%; Pred. No. 4.4e-47; Matches 139; Conservative 60; Indels 168; Gaps 2;	PI	DR N-PSDB; ACC59894.
QY	21 CPKTKCVRDAAQCGGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLM 80	PT	New human receptors and membrane-associated proteins (REMAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or infections.
Db	25 CPSECTCSRASQVECTGARIVAV-PTPLWNAMSQILNTHITELNESPFNLNISALIAR 83	PT	XX
QY	81 LSDSHISAI DPGTNDLVLKTKLTRNKSRLPRAILDKVMVLEQLDNLARLDLQN 140	PT	XX
Db	84 IENKNELSRITPGAFRNGLGSLRYSLANKLQLVLPGLFQGLDSESILLSSNOLLQIOPA 143	PT	XX
QY	141 LFOQLRNQELQINQNQLSFLPANLSSRLREKLKLLDSRNRLTHLPKGLLGAQVKLEKLL 200	PT	XX
Db	144 HFSQCSNLKELQHGNHLEYIPDGAFDHLYGLTKLNGLGNLQVLR 203	PT	XX
QY	201 LYSNQLTSDGILSNLGA LTELRLERNHLRSVAPGAFDRGNLSSLTLSGNLLESLPPA 260	PS	Claim 1; Page 187-189; 241pp; English.
Db	204 LYENRRLTDIPMGTDFGLVNLDQELAQNLQNLQYLLSPGLFHNNHNLQRLYLSNNHISQLPPS 263	XX	CC

proteins given in ABR43169 to ABR43204 and designated REMAP-1 to REMAP-36
 CC (I) have cytostatic, antiarteriosclerotic, anticonvulsant,
 CC nootropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
 CC antiinflammatory and thyromimetic activities, and can be used in gene
 CC therapy. The REMAP polypeptides and polynucleotides are useful in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of REMAP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g. AIDS,
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. They are also useful in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of REMAP. The REMAPs or their fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide
 XX

Sequence 581 AA;

Query Match 31.3%; Score 607; DB 6; Length 581;
 Best Local Similarity 37.7%; Pred. No. 4.4e-47;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
 Qy 21 CPKTKCVVRDAAQCSCGSVAHIAELGILPTNLTHILLFRMDQGILRNHSFSGMTVLQRIM 80
 Db 25 CPSECTCRASQVECTGARIAV-PTPLPNAMSLLQILNTHITELNESPFNLNISALLAR 83
 Qy 81 LSDSHISAIDPQTENDLVKLTKTIRLTKNKSRLPRAILDKATVILQFLDHNALRDLDQN 140
 Db 84 TEKNEELSRITPGAFRNLGSLRYLSLANNKLQVLPQGLFQGLDSLESLLSSNQLLQIOPA 143
 Qy 141 LFQQLRNQELGLMNQNQLSFLPANLFLSSRELKLLDSRNNLTHLPKGLLGAQVKLEKLL 200
 Db 144 HESQCSNLKELQIHNHLEYIPDGAFDHLVGLTKLNLRGNSLTHISPRVFQHLGNLQVLR 203
 Qy 201 LYSNQLTSDGILLSNLGALETLERNLRSVAPGAFDRIGNLSSLTLSGNLLSLLPPA 260
 Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLLSPGLFHNNHNNLQRLYLSNNHISQLPPS 263
 Qy 261 LFLHVSSVSRLTILENPYLELPDVLFEGMAGLRELLWLNGLTHLSTLPAAAFRNLSGLQTLG 320
 Db 264 1FMQLPQLNRLLTLEFGNSLKELSLGIIFGMPNLRELWLYDNHISLSPDNVFSNLRLQVLI 323
 Qy 321 LTRNPRLSALPGRVFGQLRERLVLALHTNALAELRDDALRGLGHLRQVSRLRNRLRALPR 380
 Db 324 LSRN-QISFISPGAFNGLTTELRELSLHTNALQDLDGNVERMLANTONISLQNNRLRQPG 382
 Qy 381 TLFRNLSSL 389
 Db 383 NIFANVNGL 391

RESULT 12
 ADC83709 ID ADC83709 standard; protein; 581 AA.
 XX AC ADC83709;
 XX DT 01-JAN-2004 (first entry)
 DE Human leucine rich repeat (LRR) Lib protein.
 XX Human; Lib; chromosome 3q29; leucine-rich repeat; LRR;
 KW transmembrane domain; adhesion; cell-cell; cell-extracellular matrix;
 KW astroglial cell; beta-amyloid processing; Alzheimer's disease;
 KW neuron death; drug screening; cell adhesion-related disease; nootropic;
 KW neuroprotective; gene therapy.
 OS Homo sapiens.
 XX PN JP2003164290-A.
 XX PD 10-JUN-2003.

RESULT 13
 ADE08191 ID ADE08191 standard; protein; 581 AA.
 XX

XX ADE08191;
 XX 29-JAN-2004 (first entry)
 XX Novel protein (useful for identifying genetic disorders) #346.
 DE novel gene; novel protein; tissue marker; molecular weight marker;
 XX chromosome marker; genetic disorder.
 XX Unidentified.
 OS
 XX WO2003054152-A2.
 XX PD 03-JTUL-2003.
 XX PF 10-DEC-2002; 2002WO-US039555.
 XX PR 10-DEC-2001; 2001US-0339739P.
 PR 11-DEC-2001; 2001US-0339453P.
 PR 14-MAR-2002; 2002US-0365091P.
 PR 14-MAR-2002; 2002US-0365384P.
 PR 12-APR-2002; 2002US-0372381P.
 PR 12-APR-2002; 2002US-0372615P.
 PR 22-APR-2002; 2002US-00128558.
 PR 24-APR-2002; 2002US-0376045P.
 PA (HYSEQ-) HYSEQ INC.
 PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QH, Wang J, Gnoth M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
 XX DR 2003-569235/53.
 XX DR N-PSDB; ADE07280.
 PT New polynucleotides, useful for expressing recombinant proteins for analysis, characterization or therapeutic use, or as markers for tissues in which the corresponding protein is preferentially expressed.
 PT
 XX PS 20; SEQ ID NO 1257; 1177pp; English.
 XX The invention comprises the amino acid and coding sequences of novel proteins. The DNA and protein sequences of the invention are useful as markers for tissues in which the corresponding protein is preferentially expressed; as molecular weight markers on gels; as chromosome markers or tags; to identify chromosomes or to map related gene positions; and to compare with endogenous DNA sequences in patients to identify potential genetic disorders. The present amino acid sequence represents a protein of the invention.
 CC Sequence 581 AA;
 SQ Query Match 31.3%; Score 607; DB 7; Length 581;
 Best Local Similarity 37.7%; Pred. No. 4.4e-47;
 Matches 139; Conservative 60; Mismatches 168; Indels 2;
 Gaps 2;
 XX DB 25 CPSECTCRASQVECTGARIYAV-PTPLPNAMSQCILNTHITELNESPFLNISALLR 83
 QY 81 LSDSHISAIIDPGTFNDLVKLFKLRLTRNKISRLPRAILDKMVLEQLFDHNLDDQN 140
 DB 84 IEXNVELSRITPGAFRNLGSRLYSLANNKLVQLVLPIGLEFOGLDSLESLLSSNQLQIOPA 143
 QY 141 LFQQQLERNLQELGLNQNQLSFLPANLFLSSRLERKLDDLSRANNLTHLPKGLGAAQVKLEXKL 200
 DB 144 HFSQCSNLKELQHGNHLEYIPDGAFDHVLGTLKLNGLKGNSLTHISPRVRFQHGNLQVLR 203
 QY 201 LYSNQLTSDSGLNSNLGAITLRLERNHLSVAFGAFDRLGNLSSLLTSGNLLS1PPA 260
 DB 204 LYENRITDIPMGTFDGLVNLQIGLQALQONQIGLSPOLFHNNHLQRLYLSNTHISQLPPS 263
 QY 261 LFLHVSSVSRRLTLFENPYLEELPDVLFGEAGLRELWLNGTHLSTLPAAAFRNLSGLQTLG 320
 DB 264 FMQLPQLNRLTLFGNSLKELSLIGIFGPMPLNRELWLVDNHISSLPDNVFSNLRLQVLI 323
 QY 321 LTRNPRLSALPRGVFOGLRELRLVLAELHTNALAELRDDALRGLGHLRQVSLRHNRLRALPR 380
 DB 324 LSRN-QISFISPGAFNGLTELRLSLLHTNALQDLDGNVFRMLANLQNISLQNNRRLQDPC 382
 QY 381 TLFRNLSSL 389
 DB 383 NIFANVNGL 391

RESULT 14
 ADN39956
 ID ADN39956 standard; protein: 581 AA.
 XX
 AC ADN39956;
 XX
 DT 17-JUN-2004 (first entry)
 XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C326.
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularistaion syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW pulmonary; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003042661-A2.
 XX PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 DR WPI; 2003-46849/44.
 XX
 PT Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO C326; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)

whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a polypeptide of the invention.

XX SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

Qy 21 CPKTKCVVRDAACGGGSVAHIAELGLPTNLTHILLFRMDQGILRNRHFSFGMTVLRQLM 80
Db 25 CPSECTCSRASQVECTGARIYAV-PTPLPNAMSILQILNTHTELNESPFLNISALIAR 83
Qy 81 LSDSHISAIDPGTFNDLVKTKLRLTRNKISRLPRAILDKMVLLQFLDHNALRDLQDN 140
Db 84 IERNELSRITPAGFRNLGSRLYSLANNKLQVLPQGLEFQGLDLESLLSSNQLLQIOPA 143
Qy 141 LFOQLRNQELGLQNQNQLSFLPANLSSRLSPNNLTHPKGLLGAQVKLEKLL 200
Db 144 HFSQCSNLKBLQLHGNHLEYIPDGAFDHLVGLTKLNKGNSLTHISPRVFOHGNLQVLR 203
Qy 201 LYSNQQLTSVDSGQLLSNLGALETIRLERNHLRSVAPGAFDRLGNLSSLTLSGNLLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLPHNNHNLQRLYLSNNHISQLPPS 263
Qy 261 LFLHVSVSRLTFLFENPYLELPDVLFGEAGLRELWLNGLTHLSTLPAAFRNLSQLQTG 320
Db 264 IFMQLPQLNRLTFLFGNSLKEIISLGLIFGPMPNLRELWLYDNHISSLPDNVFSNLRQLQVLI 323
Qy 321 LTRNPLRSALPRGVFOGLRELRLVIALHTNALAELRDLRGLGHLRQVSRLRHNRLRALPR 380
Db 324 LSRN-QISFISPQAGFNGLTELRELISLHTNALQDLDGVFRMLANLQNISLQNNRQLPQ 382
Qy 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

RESULT 15
ADN39092 ID ADN39092 standard; protein; 581 AA.
XX AC ADN39092;
XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:410.
XX KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW pulmonary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.

XX PD 22-MAY-2003.
XX PR 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0340376P.
XX PR 14-DEC-2001; 20002US-0347211P.
XX PR 08-JAN-2002; 20002US-0347349P.
XX PR 10-JAN-2002; 20002US-0355250P.
XX PR 08-FEB-2002; 20002US-0356714P.
XX PR 13-FEB-2002; 20002US-0359077P.
XX PR 20-FEB-2002; 20002US-036809P.
XX PR 29-MAR-2002; 20002US-0370110P.
XX PR 04-APR-2002; 20002US-0370110P.
XX PR 12-APR-2002; 20002US-0372246P.
XX PR 05-JUN-2002; 20002US-0386614P.
XX PR 16-JUL-2002; 20002US-0396839P.
XX PR 22-JUL-2002; 20002US-039775P.
XX PR 22-JUL-2002; 20002US-0397845P.
XX PR 09-SEP-2002; 20002US-0409450P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX PA N-PSDB; ADN39091.
XX PR WPI; 2003-468649/44.
XX DR DR
XX PS Claim 12; SEQ ID NO 410; 1385PP; English.
XX PS The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosis and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a polypeptide of the invention.
XX SQ Sequence 581 AA;

Query Match 31.3%; Score 607; DB 7; Length 581;
Best Local Similarity 37.7%; Pred. No. 4.4e-47;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
XX Qy 21 CPKTKCVVRDAACGGGSVAHIAELGLPTNLTHILLFRMDQGILRNRHFSFGMTVLRQLM 80
Db 25 CPSECTCSRASQVECTGARIYAV-PTPLPNAMSILQILNTHTELNESPFLNISALIAR 83
Qy 81 LSDSHISAIDPGTFNDLVKTKLRLTRNKISRLPRAILDKMVLLQFLDHNALRDLQDN 140
Db 84 IERNELSRITPAGFRNLGSRLYSLANNKLQVLPQGLEFQGLDLESLLSSNQLLQIOPA 143
Qy 141 LFOQLRNQELGLQNQNQLSFLPANLSSRLSPNNLTHPKGLLGAQVKLEKLL 200
Db 144 HFSQCSNLKBLQLHGNHLEYIPDGAFDHLVGLTKLNKGNSLTHISPRVFOHGNLQVLR 203
Qy 201 LYSNQQLTSVDSGQLLSNLGALETIRLERNHLRSVAPGAFDRLGNLSSLTLSGNLLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLPHNNHNLQRLYLSNNHISQLPPS 263
Qy 261 LFLHVSVSRLTFLFENPYLELPDVLFGEAGLRELWLNGLTHLSTLPAAFRNLSQLQTG 320
Db 264 IFMQLPQLNRLTFLFGNSLKEIISLGLIFGPMPNLRELWLYDNHISSLPDNVFSNLRQLQVLI 323
Qy 321 LTRNPLRSALPRGVFOGLRELRLVIALHTNALAELRDLRGLGHLRQVSRLRHNRLRALPR 380
Db 324 LSRN-QISFISPQAGFNGLTELRELISLHTNALQDLDGVFRMLANLQNISLQNNRQLPQ 382
Qy 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

Qy 684 IFOQLRNQELGLQNQLSFLPANLSSRLSPNNLTHPKGLLGAQVKLEKLL 200
Db 685 IENNELSRITPAGFRNLGSRLYSLANNKLQVLPQGLEFQGLDLESLLSSNQLLQIOPA 143
Qy 721 CPKTKCVVRDAACGGGSVAHIAELGLPTNLTHILLFRMDQGILRNRHFSFGMTVLRQLM 80
Db 725 CPSECTCSRASQVECTGARIYAV-PTPLPNAMSILQILNTHTELNESPFLNISALIAR 83
Qy 81 LSDSHISAIDPGTFNDLVKTKLRLTRNKISRLPRAILDKMVLLQFLDHNALRDLQDN 140
Db 84 IENNELSRITPAGFRNLGSRLYSLANNKLQVLPQGLEFQGLDLESLLSSNQLLQIOPA 143
Qy 141 LFOQLRNQELGLQNQLSFLPANLSSRLSPNNLTHPKGLLGAQVKLEKLL 200
Db 144 HFSQCSNLKBLQLHGNHLEYIPDGAFDHLVGLTKLNKGNSLTHISPRVFOHGNLQVLR 203

Qy 201 LYSNQLTSVDSGLLSNUGALTELRLERNHRSVAPGAFDRLGNLSSLTSGCNLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQONQIGLSPGLFHNHNHLQRLYLSNNHISQLPPS 263
Qy 261 LFLHVVSSVSRLLTILENPLEELPDVLFGEMAGLRELWNLGTHLSTLPAAAFRNLSQLQTLG 320
Db 264 IFMQLPQLNRLTLEFGNSLKELSLGIFGMPNRELWLYDNHISLSPDNVFSNLRLQQLI 323
Qy 321 LTRNPLRSALPRLGVFOGLRELRVLAHTNALAELRDDALRGLGHILRQVSLRHNRLRALPR 380
Db 324 LSRN-QISFISPRGAFNGLTTELRELSIHTNALQDLDGVFVRMLANIQNISLQDNRLRQPG 382
Qy 381 TLPRNLSSL 389
Db 383 NIFANVNGL 391

Search completed: November 29, 2004, 13:09:04
Job time : 89 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:11:04 ; Search time 93 Seconds
 (without alignments)
 1483.616 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 1938

Sequence: 1 MIRSALLSAVALLRAQFPCPKTCKCVRDAACCGSGSVAHIAELGLPTNLTHILLFRM 389

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 1575965 seqs, 354694765 residues

Total number of hits satisfying chosen parameters: 1575965

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/US06_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US07_PUB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

ALIGNMENTS

14 607 31.3 581 15 US-10-258-951-43 Sequence 43, App1
 15 607 31.3 787 15 US-10-080-334-224 Sequence 224, App
 16 607 31.3 790 14 US-10-017-161-1558 Sequence 1558, App
 17 607 31.3 790 14 US-10-292-798-1246 Sequence 1246, App
 18 607 31.3 819 14 US-10-241-220-106 Sequence 106, App
 19 607 31.3 819 17 US-10-872-972-106 Sequence 106, App
 20 583 30.1 578 15 US-10-080-334-226 Sequence 226, App
 21 541.5 27.9 391 15 US-10-080-334-227 Sequence 227, App
 22 503.5 26.0 570 14 US-10-424-233-14 Sequence 14, App1
 23 445 23.0 545 15 US-10-114-270-38 Sequence 38, App1
 24 442 22.8 545 15 US-10-312-311-20 Sequence 40, App1
 25 419.5 21.6 457 15 US-10-408-765A-632 Sequence 22, App1
 26 419.5 21.6 605 16 US-10-806-018-62 Sequence 632, App1
 27 416.5 21.5 605 16 US-10-706-791-21 Sequence 21, App1
 28 413.5 21.3 510 15 US-10-312-311-20 Sequence 20, App1
 29 413.5 21.3 605 9 US-09-782-980-62 Sequence 62, App1
 30 413.5 21.3 605 16 US-10-806-018-62 Sequence 62, App1
 31 379.5 19.6 928 16 US-10-442-658-28 Sequence 28, App1
 32 379.5 19.6 1122 16 US-10-442-658-27 Sequence 27, App1
 33 379.5 19.6 1461 16 US-10-442-658-7 Sequence 27, App1
 34 379.5 19.6 1531 10 US-09-970-944-31 Sequence 31, App1
 35 379.5 19.6 1534 10 US-09-970-944-30 Sequence 30, App1
 36 379.5 19.6 1534 16 US-10-442-658-6 Sequence 6, App1
 37 375.5 19.4 1531 10 US-09-970-944-28 Sequence 28, App1
 38 369.5 19.1 796 14 US-10-028-392-5 Sequence 5, App1
 39 369.5 19.1 1529 10 US-09-866-050A-396 Sequence 396, App1
 40 367.5 19.0 907 14 US-10-271-078-10 Sequence 10, App1
 41 366.5 18.9 1348 15 US-10-080-334-234 Sequence 234, App1
 42 366 18.9 467 15 US-10-343-348-36 Sequence 36, App1
 43 362 18.7 421 15 US-10-072-012-294 Sequence 294, App1
 44 362 18.7 1523 10 US-09-954-342-45 Sequence 45, App1
 45 361.5 18.7 1529 10 US-09-766-511B-29 Sequence 29, App1

RESULT 1

US-09-775-803-12 ; Sequence 12, Application US/09775803

; Publication No. US20030167487A1

; GENERAL INFORMATION:

; COR Therapeutics, Inc.

; APPLICANT: Ramakrishnan, Vanitha

; APPLICANT: Phillips, David

; TITLE OF INVENTION: Transgenic Animals Having a Modified Glycoprotein V

; FILE REFERENCE: 44481-504-US

; CURRENT APPLICATION NUMBER: US/09/775, 803

; CURRENT FILING DATE: 2001-02-05

; PRIORITY APPLICATION NUMBER: PCT/US99/17594

; PRIORITY FILING DATE: 1999-08-04

; PRIORITY APPLICATION NUMBER: US 60/109, 797

; PRIORITY FILING DATE: 1998-08-04

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 12

; LENGTH: 566

; TYPE: PRT

; ORGANISM: Mus musculus

; US-09-775-803-12

Query Match 100.0%; Score 1938; DB 10; Length 566;
 Best Local Similarity 100.0%; Pred. No. 4.6e-156;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLRSALLSAVALLRAQFPCPKTCKCVRDAACCGSGSVAHIAELGLPTNLTHILLFRM 60
 Db 1 MLRSALLSAVALLRAQFPCPKTCKCVRDAACCGSGSVAHIAELGLPTNLTHILLFRM 60

Qy 1 DOGILRNHFSGGMVTVLQRMLSDSHISALDPGTENDLKLTKLRLTRNKISRLPRAILDK 120
 Db 1 DOGILRNHFSGGMVTVLQRMLSDSHISALDPGTENDLKLTKLRLTRNKISRLPRAILDK 120

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	1938	100.0	566	10	US-09-775-803-12	Sequence 12, App1
2	1938	100.0	567	15	US-10-037-417-105	Sequence 105, App
3	1923	99.2	567	15	US-10-037-417-106	Sequence 106, App
4	1749	90.2	567	15	US-10-080-334-228	Sequence 228, App
5	1477	76.2	560	9	US-09-782-980-61	Sequence 61, App1
6	1477	76.2	560	10	US-09-775-803-14	Sequence 14, App1
7	1477	76.2	560	14	US-10-212-499-2	Sequence 2, App1
8	1477	76.2	560	16	US-10-806-018-61	Sequence 61, App1
9	1201	62.0	368	15	US-10-312-311-21	Sequence 21, App1
10	607	31.3	581	14	US-10-295-027-410	Sequence 410, App
11	607	31.3	581	14	US-10-295-027-1274	Sequence 1274, App
12	607	31.3	581	15	US-10-080-334-76	Sequence 76, App1
13	607	31.3	581	15	US-10-080-334-225	Sequence 225, App

QY 121 MVVLEQLFLDHNALRDQNLFOQLRNLQELGLNNQNQLSFLPANLFSSSLRELKLDLSRN 180
QY 121 MVVLEQLFLDHNALRDQNLFOQLRNLQELGLNNQNQLSFLPANLFSSSLRELKLDLSRN 180

QY 181 NLTHLPKGILLGAQVKLEKLLLYSNQLTSVDPGGLSNLGAALTEILRLERNHLSVAPGAFDR 240
QY 181 NLTHLPKGILLGAQVKLEKLLLYSNQLTSVDPGGLSNLGAALTEILRLERNHLSVAPGAFDR 240

QY 241 LGNLSSLTLSGNLLESLPPAFLHVSSVSRLTLFENPQGLRELRLVLAELRDDALR 300
QY 241 LGNLSSLTLSGNLLESLPPAFLHVSSVSRLTLFENPQGLRELRLVLAELRDDALR 300

QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFOGLRELRLVLAELRDDALR 360
QY 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFOGLRELRLVLAELRDDALR 360

QY 361 GLGHLRQVSSLRHNRRLRALPRTLFRNLSSL 389
QY 361 GLGHLRQVSSLRHNRRLRALPRTLFRNLSSL 389

Db 2 RESULT 2
US-10-037-417-105
; Sequence 105, Application US/10037417
; Publication No. US20040052806A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Patturajan, Meera
; APPLICANT: Gross, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Eddinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/318,700
 PRIOR FILING DATE: 2001-09-12
 NUMBER OF SEQ ID NOS: 227
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 105
 LENGTH: 567
 TYPE: PRT
 ORGANISM: *Mus musculus*
 US-10-037-417-105

Query	Match	Score	DB	Length
	100.0% ;	1938 ;	15 ;	567 ;
	Best Local Similarity	100.0% ;	Pred. No. 4.6e-156 ;	
Matches	389 ;	Conservative 0 ;	Mismatches 0 ;	Indels 0 ;
				Gaps 0 ;
Qy	1 MLRSAALLRAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM	60		
Db	1 MLRSAALLSAVALLRAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM	60		
Qy	61 DQGILRNHSFSGMTVILQRLMLSDSHISAI DPGTFNDLVKLKTLLRTRNKISRLPRAILDK	120		
Db	61 DQGILRNHSFSGMTVILQRLMLSDSHISAI DPGTFNDLVKLKTLLRTRNKISRLPRAILDK	120		
Qy	121 MVLLEQLFLDHNALRDLDQNLFQQLRNLQELGLNNQNQSLFLPANLFSSLRELKLLDLSRN	180		
Db	121 MVLLEQLFLDHNALRDLDQNLFQQLRNLQELGLNNQNQSLFLPANLFSSLRELKLLDLSRN	180		
Qy	241 LGNLSSLTLSGNLLESLPPALFLHVSSVSRLLTFLFENPYLEELPDVLFGE MAGLRELWLNGT	300		
Db	241 LGNLSSLTLSGNLLESLPPALFLHVSSVSRLLTFLFENPYLEELPDVLFGE MAGLRELWLNGT	300		
Qy	301 HLSTLPAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHLTNALAE LR	360		
Db	301 HLSTLPAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRELRLVLAHLTNALAE LR	360		
Qy	361 GLGHHLRQVSLRHNRRLRALPRTLFRNLSSL	389		
Db	361 GLGHHLRQVSLRHNRRLRALPRTLFRNLSSL	389		

RESULT 3
 US-10-037-417-106
 ; Sequence 106, Application US/10037417
 ; Publication No. US20040052806A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Aisobrook II, John P
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Paturajan, Meera
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Lepley, Denise M
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Li, Li
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Stone, David J
 ; APPLICANT: Boldog, Ferenc L
 ; APPLICANT: Guo, Xiaoqia
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Anderson, David W
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Miller, Charles E

Best Local Similarity 76.1%; Pred. No. 7.4e-117;
Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

1 MLRSAVLLRAQPFPCPKTCXCVRDAAQCGGSVAHIAELGLPTNLTHILLFRM 60
1 MLRGTLCAVLGLLRAQPFPCPACKCVFRAAQCGGDVARISALGLPTNLTHILLFGM 60

61 DQGILRNHSFSGMITVQLRMLSDSHISAIDPGTFNDLVKLKTLRLTRNKTSRLPRAILDK 120
61 GRGVQLQSQSPSGMTVQLRMLSDSHISAIDPGTFSDLIKLKTLRLSRNKITHLPGALLDK 120

QY 1 MVLLEQLFDHNLRDQNLQFQQLRNQQLQELQNLVNLQELNKLDDSGN 180
Db 121 MVLLEQLFDHNLRDQNLQFQQLRNQQLQELQNLVNLQELNKLDDSGN 180

QY 181 NLTHLPKGLGAAQVKLEKLLYSNQNTSVDSGLLSNLGAALTELRNHLRSVAPGAFDR 240
Db 181 NLTHLPKGLGAAQAKLERLILHNSRNLVSLDSGLLNSLGAALTELFHRNHIRSIAPGAFDR 240

QY 241 LGNLSSLTLSGNLLESLPPALFLHSSVSRLTTFENPYLEELPDVLFGEMAGLRELWNGT 300
Db 241 LPNLSSLTLSRNHLAFPLPSALFLHSHNLTLTFENPYLEELPGVLFGEMGGLQELWLNRT 300

QY 301 HLSTLPAAAFRNLSGLQTLLGTRNPRSLALPRGVFEGQGLRELRLVLAHTNALAELRDDALR 360
Db 301 QLRTLPAAAFRNLSRRLYLGTVLSPRLSALPQGAFOQLGELQVLAHSNGLTALPDGLLR 360

QY 361 GLGHLRQVSLLRNRRLALPRTLFNLSL 389
Db 361 GLGKLRLRQVSLLRNRRLALPRLFRNLSL 389

RESULT 7
US-10-212-499-2
; Sequence 2, Application US/10212499
; Publication No. US20030135036A1
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan Lewis & Bockius LLP
; STREET: 1800 M St., NW
; CITY: Washington
; STATE: DC
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/212,499
; FILING DATE: 06-Aug-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/560,814
; FILING DATE: 2000-04-28
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 1993-07-09
; APPLICATION NUMBER: US 08/195,006
; FILING DATE: 1994-02-10
; APPLICATION NUMBER: US 08/884,571
; FILING DATE: 1997-06-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Reid G. Adler
; REGISTRATION NUMBER: 30,988
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-10-212-499-2
; Query Match 76.2%; Score 1477; DB 14; Length 560;
; NUMBER OF SEQ ID NOS: 176
; Remaining Prior Application data removed - See File Wrapper or PAML.

NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 410
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-410

Query Match 31.3%; Score 607; DB 14; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
SEQ ID NO: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NOS: 1386
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-414

Query Match 31.3%; Score 607; DB 14; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;
SEQ ID NO: 1274

Qy 21 CPKTKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTVILQRLM 80
Db 25 CPSECTCSRASQVECTGARIVAV-PTPLPWNAMSLQILNTHITELNESPFLNISALIALR 83

Qy 81 LSDSHISAIIDPGTFNDLVKLTKTLLTRNKISRLPRAILDKMVILLEQFLFLDHNALRDLDQN 140
Db 84 IEKNELSRITPGAFRNLGSRLYSLANNKLIQVLPIGLFQGLDSLESLLSSNQLLQIOPA 143

Qy 141 LFOQLRNHQELGLNQNQQLSFLPANLFLSSLRELKLLDLSRNNNLTHLPKGGLGQVKLEKLL 200
Db 144 HFSQCSNLKEQLQHLGNHLEYIPDGAFDHLVGLTLKLNGLQVLR 203

Qy 201 LYSNQLTSDGGLSNLIGALTLLERNLHRSVAPGAFDRLGNLSSLTLSGNLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLSSPGLFHNNHNLQRLYLSNNNHSQQLPPS 263

Qy 261 LFLHVSSVSRLTLLFENPFLPDLFQGLREMAGLRELLWLNGLTHLSTLQVLR 320
Db 264 IFMQLPQLNRLTLLFGNSLKELSLQIGLSSPGLFHNNHNLQRLYDNHISSSLPDNVFSNLRQLQVLI 323

Qy 321 LTRNPRLSALPRGVFQGLLRELRVLAHTNALAELRDDALRGLGHLRQVSLLRHLRQVLR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDNLQVERMLNRLQQLPG 382

Qy 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

RESULT 11
US-10-295-027-1274
; Sequence 1274, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US 60/350,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14

RESULT 12
US-10-080-334-76
; Sequence 76, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Casman, Stacie J
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojaia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna T
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong

APPLICANT: Baumgartner, Jason C.
 Gerlach, Valerie K
 Spaderna, Steven K

APPLICANT: Zerhusen, Bryan D

TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 Title of Invention: Using the Same

FILE REFERENCE: 21402-275

CURRENT APPLICATION NUMBER: US/10/080,334

CURRENT FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: 60/270,523

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/322,712

PRIOR FILING DATE: 2001-09-17

PRIOR APPLICATION NUMBER: 60/311,980

PRIOR FILING DATE: 2001-08-13

PRIOR APPLICATION NUMBER: 60/330,307

PRIOR FILING DATE: 2001-10-18

PRIOR APPLICATION NUMBER: 60/278,796

PRIOR FILING DATE: 2001-03-26

PRIOR APPLICATION NUMBER: 60/281,521

PRIOR FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: 60/276,677

PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 60/311,595

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: 60/270,220

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/274,295

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: 60/318,526

PRIOR FILING DATE: 2001-09-10

PRIOR APPLICATION NUMBER: 60/286,548

PRIOR FILING DATE: 2001-04-25

PRIOR APPLICATION NUMBER: 60/291,765

PRIOR FILING DATE: 2001-05-17

PRIOR APPLICATION NUMBER: 60/270,797

PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/276,400

PRIOR FILING DATE: 2001-03-16

PRIOR APPLICATION NUMBER: 60/270,810

PRIOR FILING DATE: 2001-02-23

NUMBER OF SEQ ID NOS: 388

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO: 76

LENGTH: 581

TYPE: PRT

ORGANISM: Homo sapiens

US-10-080-334-76.

Query Match 31.3%; Score 607; DB 15; Length 581;
 Best Local Similarity 37.7%; Pred. No. 7.9e-43;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 321 LTRNPRLSALPRGVFOGLRELRLVLAHTNALAELRDDLQVSLRHNRLRALPR 380
 Db 324 LSRN-QISFISPGAFNGTLTELRELSLHTNALQDNVRQMLNRLRQLPG 382

QY 381 TLFRNLSSL 389
 Db 383 NIFANVNGL 391

RESULT 13
 US-10-080-334-225
 Sequence 225, Application US/10080334
 Publication No. US20040002584A1
 GENERAL INFORMATION:
 / APPLICANT: Pena, Carol E. A.
 / APPLICANT: Shimkets, Richard A.
 / APPLICANT: Li, Li
 / APPLICANT: Shenoy, Suresh G.
 / APPLICANT: Kekuda, Ramesh
 / APPLICANT: Spytek, Kimberly A.
 / APPLICANT: Vernet, Corine A. M.
 / APPLICANT: Malyankar, Uriel M.
 / APPLICANT: Guo, Xiaoqia
 / APPLICANT: Gusev, Vladimir Y.
 / APPLICANT: Casman, Stacie J.
 / APPLICANT: Boldog, Ferenc L.
 / APPLICANT: Furtak, Katarzyna
 / APPLICANT: Tchernev, Velizar T.
 / APPLICANT: Patturajarn, Meera
 / APPLICANT: Gangolli, Esha A.
 / APPLICANT: Padigaru, Muralidhara
 / APPLICANT: Liu, Xiachong
 / APPLICANT: Baumgartner, Jason C.
 / APPLICANT: Gerlach, Valerie
 / APPLICANT: Spaderna, Steven K.
 / APPLICANT: Zerhusen, Bryan D.

TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 Title of Invention: Using the Same

FILE REFERENCE: 21402-275

CURRENT APPLICATION NUMBER: US/10/080,334
 CURRENT FILING DATE: 2002-02-21
 PRIOR APPLICATION NUMBER: 60/270,523
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/322,712
 PRIOR FILING DATE: 2001-09-17
 PRIOR APPLICATION NUMBER: 60/311,980
 PRIOR FILING DATE: 2001-08-13
 PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/278,796
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/281,521
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/276,677
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/311,595
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/270,220
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/274,295
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: 60/318,526
 PRIOR FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/286,548
 PRIOR FILING DATE: 2001-04-25
 PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/270,797
 PRIOR FILING DATE: 2001-02-23
 NUMBER OF SEQ ID NOS: 388
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO: 76
 LENGTH: 581
 TYPE: PRT
 ORGANISM: Homo sapiens

US-10-080-334-76.

Query Match 31.3%; Score 607; DB 15; Length 581;
 Best Local Similarity 37.7%; Pred. No. 7.9e-43;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

QY 21 CPKTCCKCYRDAAQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTVLQRLM 80
 Db 25 CPSECTCSRASQVECTGARIYAV-PTPLPNAMSILQILNTHITELNESPFNLNISALALR 83

QY 81 LSDSHISAIIDPGTFNDLVKLFTRLTRNKKISRLPRAILDKMVLEQLFDHNLARLDQN 140
 Db 84 IKNBLSRITPGAFRNLGSRLYSLANKNLQVLPIGLFOGLDSLESLLSSNOLLQIOPA 143

QY 141 LFQQLRLNQELGLNQNQNLSSFLPANLFSSSLRLKLLDSRNNLTHLPKGLGAQVKLEKLL 200
 Db 144 HFSQCSNLKELQHGNHLEYIPDGAFDHLVGLTKMLGKNSLTHISPRVVFQHGNLQVLR 203

QY 201 LYSNOLTSVDSGLSNNLGAALTELRNHLRSVAPGAFDRLGNLSSLTLSGNLLESLPPA 260
 Db 204 LYENRLTDIPMGTFDGLVNLQCNQIGLSPGLFHNNHLQRLYLSNNHSQLPPS 263

QY 261 LFLHVSSVSRLTFLFENPLEELPDVLFGEMAGLRELWLNGLTHLSTLPAAFRNLSQLTLG 320
 Db 264 IFNQLPQLNRLTFLFGNSLKELSLGIFGMPNLRDNVFSNLRLQQLVLI 323

PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: 60/286,548
 PRIOR FILING DATE: 2001-02-21
 PRIOR APPLICATION NUMBER: 60/270,797
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/276,400
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/311,595
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/270,220
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/311,980
 PRIOR FILING DATE: 2001-08-13
 PRIOR APPLICATION NUMBER: 60/291,765
 PRIOR FILING DATE: 2001-09-17
 PRIOR APPLICATION NUMBER: 60/270,797
 PRIOR FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: 60/276,400
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/270,810

```

; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 225
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-080-334-225

Query Match 31.3%
Best Local Similarity 37.7%
Matches 139; Conservative

 21 CPKTCKCVVRDAQCSG
 25 CPSECTCSRASQVECTG

 81 LSDSHISAIIDPGTFNDL
 84 IEKNELSRTPGAFRNLQ

 141 LFQQQLRNLQELGLNQNQI
 144 HFSQCSMNLKELOLHGNH

 201 LYSNQQLTSVDSGLLSNL
 204 LYENRRLTDIPMGTFDGLY

 261 LFLHVSSVSRLTLFENPI
 264 IFMQLPQLNRLTLFGNSI

 321 LTRNPRLSALPRGVFOQI
 324 LSRN-QISFISPGAFNGI

 381 TLFRNLSSL 389
 383 NIFANVNGL 391

RESULT 14
US-10-258-951-43
Sequence 43, Application US/10
Publication No. US20040033504A
GENERAL INFORMATION:
APPLICANT: Agarwal, Pankaj
APPLICANT: Murdock, Paul R.
APPLICANT: Rizvi, Safia K.
APPLICANT: Smith, Randall F.
APPLICANT: Xiang, Zhaoying
APPLICANT: Kabnick, Karen
APPLICANT: Lai, Ying-Ta
APPLICANT: Xie, Qing
TITLE OF INVENTION: NOVEL COM
FILE REFERENCE: GP50025
CURRENT APPLICATION NUMBER: US/10
CURRENT FILING DATE: 2002-10
PRIOR APPLICATION NUMBER: PCT
PRIOR FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: 60/
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: 60/
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSEQ for Windows
SEQ ID NO 43
LENGTH: 581
TYPE: PRT

```

ORGANISM: Homo sapiens
US-10-258-951-43

Query Match 31.3%; Score 607; DB 15; Length 581;
Best Local Similarity 37.7%; Pred. No. 7.9e-43;
Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2

Qy 21 CPKTCKCVVRDAACQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTVLUQRIM 80
Db 25 CPSECTCSRASQVECTGARIHAV-PTPLPWNAMSILQILNTHITELNESPFLNISALIALR 83

Qy 81 LSDSHISALDPGTFNNDLVKLKTRLTRNKISRLPRAILDKMVLLEQLFDHNAIRDLDQN 140
Db 84 IEKNELSRITPGAFRNLGSLRYLSANNKLQVLPIGLFOQLDSLESLLSSNQLLQIQPA 143

Qy 141 LFQQILRNQELGLNQNQQLSFLPANLPSSSLREIKLDDLSRNNNLTHLPKGILLGAQVKLEKIL 200
Db 144 HFSQCSNLKELQLHGNHLEYIPDGAFDHLVGLITKLNLKNSLTHISPRVFQHLCGNLQVLR 203

Qy 201 LYSNQLTSDSGLLSNLGALTEILRLERNHLLRSVAPGAFDRLGNLSSLTLSGNLLESLPPA 260
Db 204 LYENRLTDIPMGTFDGLVNLQELALQQNQIGLSPGLFHNHHNLQRLYLSNNHISQLPPS 263

Qy 261 LFLHVSSVSRLTTFENPYLEELPFDVLFEGEMAGLRELWLNGTHLSTLPAAAFRNLSGLQTLG 320
Db 264 IFMQLPOLNRLTTFGNSLKEISLGIFGPMNLRRELWLVDNHNHISSLPDNVFSNLRLQVLI 323

Qy 321 LTRNPRLSALPGRGVFGQLRELRLVLAELRDDALRGLGHLRQVSLLRHNRLRALPR 380
Db 324 LSRN-QISFISPGAFNGLTELRELSLHTNALQDLDGNTVFRMLANLQNISSLQNNRRLRQLPG 382

Qy 381 TLFRNLSSL 389
Db 383 NIFANVNGL 391

RESULT 15
US-10-080-334-224
Sequence 224, Application US/10080334
Publication No. US20040002584A1

GENERAL INFORMATION:

APPLICANT: Pena, Carol E. A.
APPLICANT: Shimkets, Richard A.
APPLICANT: Li, Li
APPLICANT: Shenoy, Suresh G
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A.
APPLICANT: Vernet, Corine A. M.
APPLICANT: Malyankar, Uriel M
APPLICANT: Guo, Xiaoqiao
APPLICANT: Gusev, Vladimir Y
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Furtak, Katarzyna
APPLICANT: Tchernev, Velizar T
APPLICANT: Patturajan, Meera
APPLICANT: Gangolli, Esha A
APPLICANT: Padigaru, Muralidhara
APPLICANT: Liu, Xiaohong
APPLICANT: Baumgartner, Jason C.
APPLICANT: Gerlach, Valerie
APPLICANT: Spaderna, Steven K
APPLICANT: Zerhusen, Bryan D

APPLICANT: TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of Using the Same
FILE REFERENCE: 21402-275

CURRENT APPLICATION NUMBER: US/10/080,334
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/270,523
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/322,712
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/311,980

```

; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 224
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-224

```

Query Match 31.3%; Score 607; DB 15; Length 787;
 Best Local Similarity 37.7%; Pred. No. 1.2e-42;
 Matches 139; Conservative 60; Mismatches 168; Indels 2; Gaps 2;

Qy 21 CPKTKCVRDAAQCGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMVTVLQRLM 80
 Db 231 CPSECTCSRASQVECTGARIVAV-PTPLEWNAAMSLLQILNTHITELNESPPFLNISALIALR 289
 Qy 81 LSDSHISAIIDPGTFNDLVKLTKLRLTRNKISRLPRAILDKMVLLIEQLFDNHNLRDQN 140
 Db 290 IEGNELSRITPGAFRNGLSRYLSLANNKLQVLPPIGLFQGLDSLESLLSSNQLLQIOPA 349
 Qy 141 LFQQLRNQELGLNQNQLSFLPANLSSLRELKLUJDLSRNNNTLHKGLGQVKLEKLL 200
 Db 350 HFSQCSNKLQELQHGNHLEYIFDGAFDHVLGVLTKLNLRKNSLTHISPRVFOHGLNLQVLR 409
 Qy 201 LYSNQLTSDSGLLSNLGALETLLERNLRSVAPGAFDRLGNLSSLTSGNLLESLPPA 260
 Db 410 LYENRLTDIPMGTFDGLVNLQELALQONQIGLLSPGLFHNNHNLQRLYLSNNHISQLPPS 469
 Qy 261 LFLHVSYSRRLTIFENPYLELPDVLFGEMAGLRELLWNGTHLSTLPAAAFRNLSGLQTLG 320
 Db 470 IFMQLPQLNRLTIFGNSLKELSLGLGIFGEMPNLRELLWLYDNHISLFDNVFSNLRQQLI 529
 Qy 321 LTRNPRLSALPRLGVOGLRELLRVLAELHTNALAELRDDALRGLRQVSLRNRLRALPR 380
 Db 530 LSRN-QISEFISPCAFNGLITEIRELISLTNALQDLDGVVERMLANLQNIISQNNRRLQRLPG 588
 Qy 381 TLFRNLSSL 389
 Db 589 NIFANVNL 597

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 12:59:28 ; Search time 109 Seconds
 (without alignments)

2053.401 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389
 Perfect score: 1938
 Sequence: 1 MRLSALLSAVLLRAQPFPP.....IRHNRLRALPRTLFRNLSL 389

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : UniProt_02:
 1: uniprot_sprot:
 2: uniprot_trembl:
 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1938	100.0	567	2	Q9QZU3		Q9qzus3 mus musculu
2	1923	99.2	567	1	GPV_MOUSE		Q08742 mus musculu
3	1749	90.2	567	1	GPV_RAT		Q08770 rattus norv
4	1477	76.2	560	1	GPV_HUMAN		P40197 homo sapien
5	782	40.4	637	2	Q6DCV7		Q6dcv7 xenopus lae
6	607	31.3	581	1	JR15_HUMAN		Q8tf66 homo sapien
7	607	31.3	581	2	Q7RTN7		Q7rtn7 homo sapien
8	600	31.0	579	2	Q80X72		Q80x72 mus musculu
9	583	30.1	578	1	JR15_RAT		Q8r5m3 rattus norv
10	541.5	27.9	391	2	Q9D3K0		Q9d3k0 m mus musculu
11	503.5	26.0	562	2	Q8R113		Q8r113 mus musculu
12	503.5	26.0	570	2	Q9DBB9		Q9dbb9 mus musculu
13	445	23.0	557	2	Q8N5V4		Q8n5v4 homo sapien
14	445	23.0	560	2	Q86SU4		Q86su4 homo sapien
15	419.5	21.6	605	1	ALS_HUMAN		P35858 homo sapien
16	416.5	21.5	536	1	CBP8_HUMAN		P22792 homo sapien
17	416.5	21.5	605	2	Q8TAIO		Q8tay0 homo sapien
18	413.5	21.3	605	1	ALS_PAPHA		Q02833 papio hamad
19	397	20.5	3127	2	Q7PTD4		Q7ptd4 anopheles g
20	384	19.8	1000	2	Q7QF88		Q7qf88 anopheles g
21	382.5	19.7	603	1	ALS_RAT		P35859 rattus norv
22	382	19.7	603	1	ALS_MOUSE		P70389 mus musculu
23	382	19.7	603	2	AAF69482		Aaf69482 mus musculu
24	382	19.7	687	2	Q9JIL0		Q9jil0 mus musculu
25	381.5	19.7	603	2	Q70211		O70211 rattus norv
26	379.5	19.6	1531	1	SLT1_RAT		O88279 rattus norv
27	379.5	19.6	1534	1	SLT1_HUMAN		O75093 homo sapien
28	379	19.6	417	2	Q6EB4J7		Q6ej4j7 petromyzon
29	375.5	19.4	1531	1	SLT1_MOUSE		Q80tr4 mus musculu
30	372.5	19.2	1058	2	Q76FN7		Q76fn7 tachypyleus
31	372.5	19.2	1058	2	BAD12073		Bad12073 tachypyleu

32	370	19.1	890	2	Q7Q941	anopheles g
33	369.5	19.1	766	1	SLT2_RAT	Q9wvc1 rattus norv
34	367.5	19.0	907	1	LGR5_MOUSE	Q9z1P4 mus musculu
35	367.5	19.0	1521	1	SLT2_MOUSE	Q9rlb9 mus musculu
36	366.5	18.9	1392	2	Q9VAD1	Q9vad1 drosophila
37	366	18.9	1523	1	SLT3_RAT	Q88280 rattus norv
38	365	18.8	451	2	Q7QIS1	Q7qis1 anopheles g
39	364.5	18.8	615	2	Q9VZ84	Q9vz84 drosophila
40	364.5	18.8	1529	2	Q7ZX12	Q7zxi2 xenopus lae
41	362	18.7	1523	1	SLT3_MOUSE	Q9wvb4 mus musculu
42	361.5	18.7	1529	1	SLT2_HUMAN	Q94813 homo sapien
43	361.5	18.7	1530	2	Q90WZ3	Q90wz3 xenopus lae
44	361	18.6	1173	2	Q7QHH1	Q7qhh1 anopheles g
45	360	18.6	737	2	Q965M3	Q965m3 caenorhabdi

ALIGNMENTS

RESULT 1							
Q9QZU3		PRELIMINARY;		PRT;		567 AA.	
ID	Q9QZU3;	AC					
AC		DT	01-MAY-2000 (TREMBLrel. 13, Created)				
		DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
		DT	01-OCT-2004 (TREMBLrel. 28, Last annotation update)				
		DE	Platelet glycoprotein V (Fragment) (Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone: 6030400G03 (Product:glycoprotein 5 (platelet), full insert sequence)				
		DE	(Fragment)				
		GN	Name=Gp5;				
		OS	Mus musculus (Mouse)				
		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
		OC					
		NCBI_TaxID	=10090;				
		[1]					
		SEQUENCE FROM N.A.					
		RP	STRAIN=129/SV;				
		RC					
		RX	MDLINE=20027548; PubMed=10557321;				
		RA	Ramakrishnan V.; Reeves P.S.; DeGuzman F.; Deshpande U.,				
		RA	Ministri-Madrid K.; DuBridge R.B.; Phillips D.R.;				
		RT	"Increased thrombin responsiveness in platelets from mice lacking glycoprotein V";				
		RL	Proc. Natl. Acad. Sci. U.S.A. 96:133336-13341 (1999).				
		RP	SEQUENCE FROM N.A.				
		RC	STRAIN=C57BL/6J; TISSUE=Testis;				
		RX	MDLINE=99279253; PubMed=10349636;				
		RA	Carninci P.; Hayashizaki Y.;				
		RA	"High-efficiency full-length cDNA cloning.";				
		RT	RT				
		RL	High-efficiency full-length cDNA cloning.";				
		RL	RT				
		RL	Meth. Enzymol. 303:19-44 (1999).				
		[3]					
		SEQUENCE FROM N.A.					
		RP	STRAIN=C57BL/6J; TISSUE=Testis;				
		RC	MDLINE=21085660; PubMed=11217851;				
		RA	RIKEN FANTOM Consortium;				
		RT	"Functional annotation of a full-length mouse cDNA collection."				
		RL	Nature 409:685-690 (2001).				
		RN					
		SEQUENCE FROM N.A.					
		RP	STRAIN=C57BL/6J; TISSUE=Testis;				
		RC	The FANTOM Consortium,				
		RA	RA				
		RA	"Analysis of the mouse transcriptome based on functional annotation of the mouse transcriptome."				
		RT	RT				
		60,770	full-length cDNAs."				
		RT	RT				
		RN	Nature 420:563-573 (2002).				
		[5]					
		SEQUENCE FROM N.A.					
		RP	STRAIN=C57BL/6J; TISSUE=Testis;				
		RC	MDLINE=20499374; PubMed=11042159;				
		RA	Carninci P.; Shibusawa Y.; Hayashiz				

critical initiating event in hemostasis (By similarity).

REPEAT 314 LRR 11. SUBCELLULAR LOCATION: Type I membrane protein.

REPEAT 338 361 -.- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.

REPEAT 362 385

REPEAT 386 409

CARBOHYD 51 51 N-linked (GICNAC. . .) (Potential).

CARBOHYD 67 67 N-linked (GICNAC. . .) (Potential).

CARBOHYD 181 181 N-linked (GICNAC. . .) (Potential).

CARBOHYD 243 243 N-linked (GICNAC. . .) (Potential).

CARBOHYD 298 298 N-linked (GICNAC. . .) (Potential).

CARBOHYD 312 312 N-linked (GICNAC. . .) (Potential).

CARBOHYD 385 385 N-linked (GICNAC. . .) (Potential).

SEQUENCE 567 AA; 63467 MW; C48643AA73967AD CRC64;

Query Match 99.2%; Score 1923; DB 1; Length 567;

Best Local Similarity 99.2%; Pred. No. 2.4e-121;

Matches 386; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DR 1 MLRSAVLLRAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60

DR 1 MERSAVLAVPLLAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60

DR 61 DQGILRNHSFSGMTVQLRMLSDSHISAIIDPGTENDLVKLKTLRTRNKISRLPRAILDK 120

DR 61 DQGILRNHSFSGMTVQLRQMLSDSHISAIIDPGTENDLVKLKTLRTRNKISRLPRAILDK 120

DR 121 MVLLEQLFLDHNALRDLDDONLFOQLRNQELGLNNQNQLSFLPANLFESSLREKLRLDDLSRN 180

DR 121 MVLLEQLFLDHNALRDLDDONLFOQLRNQELGLNNQNQLSFLPANLFESSLREKLRLDDLSRN 180

DR 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDR 240

DR 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDR 240

DR 241 LGNLSSLTLSGNLLESLPPALFLHVSSVSRLTTFENPLEELPDVLFGE MAGLRELWLNGT 300

DR 241 LGNLSSLTLSGNLLESLPPALFLHVSSVSRLTTFENPLEELPDVLFGE MAGLRELWLNGT 300

DR 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRRELVLLAELRDDALR 360

DR 301 HLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRRELVLLAELRDDALR 360

DR 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389

DR 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSSL 389

DR 387 409

DR 51 51 N-linked (G1cNAC. . .) (Potential).

DR 181 181 N-linked (G1cNAC. . .) (Potential).

DR 243 243 N-linked (G1cNAC. . .) (Potential).

DR 298 298 N-linked (G1cNAC. . .) (Potential).

DR 312 312 N-linked (G1cNAC. . .) (Potential).

DR 385 385 N-linked (G1cNAC. . .) (Potential).

DR 498 498 N-linked (G1cNAC. . .) (Potential).

SQ 567 AA; 63344 MW; CA10708E0D03707F CRC64;

Query Match 90.2%; Score 1749; DB 1; Length 567;

Best Local Similarity 89.2%; Pred. No. 1.3e-109;

Matches 347; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

DR 1 MLRSAVLLRAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60

DR 1 MRSVLLSAVLSVGAQPFPCPKTCKCVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60

QY 61 DQGILRNHSFSGMTVQLRMLSDSHISAIIDPGTENDLVKLKTLRTRNKISRLPRAILDK 120

Db 61 DRGVLOSHSFSGMTVQLRMLSDSHISAIIDPGTENDLVKLKTLRTRNKISRLPRAILDK 120

QY 121 MVILLEQLFHDHNALRDLDDONLFOQLRNQELGLNNQNQLSFLPANLFESSLREKLRLDDLSRN 180

Db 121 MVILLEQLFHDHNALRDLDDONLFOQLRNQELGLNNQNQLSFLPANLFESSLREKLRLDDLSRN 180

QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDR 240

QY 181 NLTHLPKGLLGAQVKLEKLLYSNQLTSVDSGLLSNLGALTELRLERNHLRSVAPGAFDR 240

-.- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; Z69594; CAA93440.1; -.

CC HSSP; Q9BZER6; 1OZN.

CC RGD; 2724; Gp5.

CC InterPro; IPR001611; LRR.

CC InterPro; IPR000483; LRR_Cterm.

CC InterPro; IPR000372; LRR_Nterm.

CC InterPro; IPR003591; LRR_typ.

CC Pfam; PF00560; LRR; 14.

CC PRINTS; PRO0019; LEURICHRPT.

CC SMART; SM00082; LRRCT; 1.

CC SMART; SM00013; LRRNT; 1.

CC SMART; SM00369; LRR_TYP; 10.

CC Blood coagulation; Cell adhesion; Glycoprotein; Leucine-rich repeat; Platelet; Repeat; Signal; Transmembrane.

CC SIGNAL 1 16 Potential.

CC CHAIN 17 567 Platelet glycoprotein V.

CC DOMAIN 17 522 Extracellular (Potential).

CC TRANSMEM 523 543 Potential.

CC DOMAIN 544 567 Cytoplasmic (Potential).

CC REPEAT 73 96 LRR 1.

CC REPEAT 97 120 LRR 2.

CC REPEAT 122 144 LRR 3.

CC REPEAT 145 168 LRR 4.

CC REPEAT 169 192 LRR 5.

CC REPEAT 194 216 LRR 6.

CC REPEAT 217 240 LRR 7.

CC REPEAT 241 264 LRR 8.

CC REPEAT 266 288 LRR 9.

CC REPEAT 289 312 LRR 10.

CC REPEAT 314 337 LRR 11.

CC REPEAT 338 361 LRR 12.

CC REPEAT 362 385 LRR 13.

CC REPEAT 387 409 LRR 14.

CC CARBOHYD 51 51 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 181 181 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 243 243 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 298 298 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 312 312 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 385 385 N-linked (G1cNAC. . .) (Potential).

CC CARBOHYD 498 498 N-linked (G1cNAC. . .) (Potential).

SQ SEQUENCE 567 AA; 63344 MW; CA10708E0D03707F CRC64;

RESULT 3

GPV_RAT

STRAIN=GPV_RAT

SEQUENCE FROM N.A.

O08770; Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Buteleostomi; Murinae; Rattus; NCBITaxID=10116;

15-JUL-1998 (Rel. 36, Created) 15-JUL-1998 (Rel. 36, Last sequence update) 05-JUL-2004 (Rel. 44, Last annotation update)

05-JUL-2004 (Rel. 44, Last annotation update)

Platelet glycoprotein V precursor (GPV) (CD42D). Name=Gp5;

Rattus norvegicus (Rat). Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus; NCBITaxID=10116;

SEQUENCE FROM N.A.

MDLINE=97275136; PubMed=9129030; Ravanat C., Morales M., Azorsa D.O., Moog S., Schuhler S., Grunert P., Loew D., van Dorsselaer A., Cazenave J.-P., Lanza F.; "Gene cloning of rat and mouse platelet glycoprotein V: identification of megakaryocyte-specific promoters and demonstration of functional thrombin cleavage." Blood 89:3253-3262 (1997).

-.- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand factor receptor and mediates von Willebrand factor-dependent platelet adhesion to blood vessels. The adhesion of platelets to injured vascular surfaces in the arterial circulation is a

241 LGNLSSLTSGNILESLPPAFLFHSSVSRLLTFENPLEELPDVLFGEAGLRELWLNGT 300
 241 LGNLSSLTSGNILESLPPAFLFHSSVSRLLTFENPLEELPDVLFGEAGLRELWLNGT 300

QY 301 HLSTPLAAFRNLSSGLQTLGLTRNPRSLALPRGVFOGLRELVLALHTNALAELRDDALR 360
 DB 301 HLRTPLAAFRNLSSGLQTLGLTRNPLLSALPPGMFHGLTELRLVLAHTNALEELPEDALR 360

QY 361 GLGHLRQVSLRHNRLPRLPRTLFRNLSSL 389
 DB 361 GLGRLRQVSLRHNRLPRLPRTLFRNLSSL 389

RESULT 4
 GPV_HUMAN
 ID GPV_HUMAN STANDARD; PRT; 560 AA.
 AC P40197;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Platelet glycoprotein V precursor (GPV) (CD42D).
 GN Name=GP5;
 OS Homo sapiens (Human).
 OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=93391348; PubMed=7690959;
 RA Hickey M.J., Hagen F.S., Yagi M., Roth G.J.;
 RT "Human platelet glycoprotein V: characterization of the polypeptide
 and the related Ib-V-IX receptor system of adhesive, leucine-rich
 glycoproteins";
 RT glycoprotein V. A member of the leucine-rich glycoprotein family
 cleaved during thrombin-induced platelet activation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8327-8331(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Platelet;
 RX MEDLINE=94012616; PubMed=8407908;
 RA Lanza F., Morales M., de la Salle C., Cazenave J.-P., Clemetson K.J.,
 RA Shimomura T., Phillips D.R.;
 RT "Cloning and characterization of the gene encoding the human platelet
 glycoprotein V. A member of the leucine-rich glycoprotein family
 cleaved during thrombin-induced platelet activation.";
 RL J. Biol. Chem. 268:20801-20807(1993).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Platelet;
 RX MEDLINE=90275263; PubMed=2350580;
 RA Shimomura T., Fujimura K., Maehama S., Takemoto M., Oda K.,
 RA Fujimoto T., Suzuki M., Ichihara-Tanaka K., Titani K.,
 RA Kuramoto A.;
 RT "Rapid purification and characterization of human platelet
 glycoprotein V: the amino acid sequence contains leucine-rich
 repetitive modules as in glycoprotein Ib.";
 RL Blood 75:2349-2356 (1990).
 RN [4]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Platelet;
 RX MEDLINE=90321220; PubMed=2372284;
 RA Roth G.J., Church T.A., McMullen B.A., Williams S.A.;
 RT "Human platelet glycoprotein V: a surface leucine-rich glycoprotein
 related to adhesion.";
 RT Biochem. Biophys. Res. Commun. 170:153-161 (1990).
 CC -!- FUNCTION: The GPIb-V-IX complex functions as the von Willebrand
 CC factor receptor and mediates von Willebrand factor-dependent
 CC platelet adhesion to blood vessels. The adhesion of platelets to
 CC injured vascular surfaces in the arterial circulation is a
 CC critical initiating event in hemostasis.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Platelets and megakaryocytes.
 CC -!- PTM: The N-terminus is blocked.

CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL; L11238; AAA03069.1;
 CC DR EMBL; Z23091; CAA80637.1;
 CC DR PIR; A48030; A60164.
 CC DR HSSP; P07359; 1MOZ.
 CC DR Genew; HGNC:4443; GPS.
 CC DR MIM; 173511;
 CC DR GO:0005887; C:integral to plasma membrane; TAS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 14.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 KW Blood coagulation; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Leucine-rich repeat; Platelet; Repeat; Signal;
 KW Transmembrane.
 FT SIGNAL 1 16
 FT CHAIN 17 560
 FT DOMAIN 17 523
 FT TRANSMEM 524 544
 FT DOMAIN 545 560
 FT REPEAT 73 96
 FT REPEAT 97 120
 FT REPEAT 122 144
 FT REPEAT 145 168
 FT REPEAT 169 192
 FT REPEAT 194 216
 FT REPEAT 217 240
 FT REPEAT 241 264
 FT REPEAT 266 288
 FT REPEAT 289 312
 FT REPEAT 314 337
 FT REPEAT 338 361
 FT REPEAT 362 385
 FT REPEAT 386 409
 FT CARBOHYD 51 51
 FT CARBOHYD 181 181
 FT CARBOHYD 243 243
 FT CARBOHYD 267 267
 FT CARBOHYD 298 298
 FT CARBOHYD 312 312
 FT CARBOHYD 385 385
 FT CARBOHYD 499 499
 FT CONFLICT 73 74
 FT CONFLICT 109 109
 FT CONFLICT 130 130
 FT CONFLICT 136 138
 FT CONFLICT 209 209
 FT CONFLICT 267 267
 FT CONFLICT 327 327
 FT CONFLICT 478 478
 FT CONFLICT 509 509
 FT CONFLICT 560 AA; 60959 MW; B1CDB04AF8AF7115 CRC64;

Query Match 76.2%; Score 1477; DB 1; Length 560;
 Best Local Similarity 76.1%; Pred. No. 2.7e-91;
 Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;
 1 MIRSALLSAVLLRAQFPFCPKTCKVVRDAAQCSGGSVAHIAELGLPTNLTHILLFRM 60
 1 MLRGTLCAVGLGLRAQFPFCPEACKCVFRDAAQCSGGDVARTSALGLPTNLTHILLFRM 60
 61 DQGILRNHSSSGMTVLQRLMLSDSHISAIIDPGTFNDLVKLTLLTRNKISRLPRAILDK 120

Qy
 Db
 Qy

Db	61	GRGVHQSQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKTLRSLRNKITHLPGALLDK	120	RA	Klein S., Strausberg R.;	
	121	MVLLQLFDNHNALRDLIDQNLFOQLRNQNQLSFLPANLFSSLRELKLLDSRN	180	RL	Submitted (JUL-2004) to the EMBL/GenBank/DDJB databases.	
QY	121	MVLLQLFDNHNALRGLDQNMFQLVNLHSNRLLHSNRLVSLDSGLNLSGLALTELOFHRNHIRSIAPGAFDR	240	DR	EMBL; BC077882; AAH77882.1; -	
Db	181	NLTHILPKGLLGAQVKLEKKLLYSNQLTSVDSGLLSNLIGALTELRLERNHRSVAPGAFDR	240	KW	Hypothetical protein.	
QY	181	NLTHILPKGLLGAQAKLERLLHSNRLVSLDSGLNLSGLALTELOFHRNHIRSIAPGAFDR	240	SQ	SEQUENCE 637 AA; 72764 MW; 2D174E8895789F23 CRC64;	
		Query Match 40.4%; Score 782; DB 2; Length 637;				
		Best Local Similarity 45.3%; Pred. No. 2.3e-44;				
		Matches 167; Conservative 58; Mismatches 144; Indels 0; Gaps 0;				
QY	21	CPRTCKCVVRDAACQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTVLQRLM	80	QY	21 CPRTCKCVVRDAACQCSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTVLQRLM	80
Db	18	CPTLCTCRLLKDAVFCQGSPSIKDIGSLLPSNFTYIHTINTLATEITDKSFGNMPITLRLR	77	Db	18 CPTLCTCRLLKDAVFCQGSPSIKDIGSLLPSNFTYIHTINTLATEITDKSFGNMPITLRLR	77
QY	81	LSDSHISAIIDPGTFNDLVKLKTRLTRNKISRLRAILDKMVILLEQLFLDHNALRDLQDN	140	QY	81 LSDSHISAIIDPGTFNDLVKLKTRLTRNKISRLRAILDKMVILLEQLFLDHNALRDLQDN	140
Db	78	LEDSRLTFITRDAFKSLPQLKSLKLTNNKLETLPAGVFDSLFFYLEQIFIGVNHLSSLHPN	137	Db	78 LEDSRLTFITRDAFKSLPQLKSLKLTNNKLETLPAGVFDSLFFYLEQIFIGVNHLSSLHPN	137
QY	141	LFQQLRNQELGLNQNQLSFLPANLFSSLRELKLLDSRNLTALPDGLLR	200	QY	141 LFQQLRNQELGLNQNQLSFLPANLFSSLRELKLLDSRNLTALPDGLLR	200
Db	138	LFCCLQHLKELILRNQNQLTSLPNELLRNLTTELITLNLSRNKISHLPVSIFSSLTKLKKLH	197	Db	138 LFCCLQHLKELILRNQNQLTSLPNELLRNLTTELITLNLSRNKISHLPVSIFSSLTKLKKLH	197
QY	201	LYSNQLTSDSGLISNLGALTELRNHLRSVAPGAFDRLGNLSSLTLSGNLESLPPA	260	QY	201 LYSNQLTSDSGLISNLGALTELRNHLRSVAPGAFDRLGNLSSLTLSGNLESLPPA	260
Db	198	LYENQLLTITSSAFNNLGEELALAYNSNISQSIAPDAFHHLPKLRLLNLSKVNKLHFLPYG	257	Db	198 LYENQLLTITSSAFNNLGEELALAYNSNISQSIAPDAFHHLPKLRLLNLSKVNKLHFLPYG	257
QY	261	LFLHVSSVSRRLTFENPLEELPDVLFGEMAGLRELWLNGLSTLPAAAFRNLSGLQTLG	320	QY	261 LFLHVSSVSRRLTFENPLEELPDVLFGEMAGLRELWLNGLSTLPAAAFRNLSGLQTLG	320
Db	258	LFLHLPQLSVLTLYDNPPLKELPDVIFGKMEMNLTSWLWYDTHLATIPNFVFCNLTLQLLV	317	Db	258 LFLHLPQLSVLTLYDNPPLKELPDVIFGKMEMNLTSWLWYDTHLATIPNFVFCNLTLQLLV	317
QY	321	LTRNPRLSALPRGVFQGLRQLRQVSLRHNRLRALPR	380	QY	321 LTRNPRLSALPRGVFQGLRQLRQVSLRHNRLRALPR	380
Db	318	LTRNPQLESLPADAFSGLSNLLELSLHSNNLSSIDQDLFQNLQQLEKLSLYSNNLKVULSE	377	Db	318 LTRNPQLESLPADAFSGLSNLLELSLHSNNLSSIDQDLFQNLQQLEKLSLYSNNLKVULSE	377
QY	381	TLFRNLISSL	389	QY	381 TLFRNLISSL	389
Db	378	NMFYNLSNL	386	Db	378 NMFYNLSNL	386
		RESULT 6				
		LRI5_HUMAN				
		ID LR15_HUMAN	STANDARD;			
		AC Q8TF66;				
		DT 10-OCT-2003 (Rel. 42, Created)				
		DT 10-OCT-2003 (Rel. 42, Last sequence update)				
		DT 05-JUL-2004 (Rel. 44, Last annotation update)				
		DE Leucine-rich repeat-containing protein 15 precursor (hLib).				
		GN Name=LRRC15; Synonyms=LIB;				
		OS Homo sapiens (Human)				
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
		NCBI_TAXID=9606;				
		RN [1]				
		SEQUENCE FROM N.A.				
		RC TISSUE=Spleen;				
		RC MEDLINE=22341132; PubMed=12454917;				
		RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,				
		RA Richardson P.; RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative."				
		RT Dev. Dyn. 225:384-391 (2002).				
		RN [2]				
		RP SEQUENCE FROM N.A.				
		RC TISSUE=Spleen;				
		RC PubMed=12477932;				
		RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
		RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
		RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
		RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
		RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
		RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
		RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
		RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,				
		RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
		RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
		RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
		RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
		RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
		RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
		RA Rodriguez A.C., Grimwood J., Myers R.M., Butterfield Y.S.,				
		RA Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E.,				
		RA Jones S.J., Marras M.A.; RT "Generation and initial analysis of more than 15,000 full-length human				
		RT and mouse cDNA sequences.";				
		RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
		[3]				
		SEQUENCE FROM N.A.				
		TISSUE=Spleen;				
		RP				
		RC				
		CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
		CC between the Swiss Institute of Bioinformatics and the EMBL outstation -				
		CC the European Bioinformatics Institute. There are no restrictions on its				
		CC use by non-profit institutions as long as its content is in no way				
		CC modified and this statement is not removed. Usage by and for commercial				
		CC entities requires a license agreement (See http://www.isb-sib.ch/announce/				
		CC or send an email to license@isb-sib.ch).				
		CC DR				

DR	HSSP; Q9BZR6; IOZN.	PRELIMINARY;	PRT;	581 AA.
DR	Genew; HGNC:20818; LRRC15.			
DR	InterPro; IPR011061; Antihaemostatic.			
DR	InterPro; IPR01611; IRR.			
DR	InterPro; IPR00483; IRR_Cterm.			
DR	InterPro; IPR00372; IRR_Nterm.			
DR	InterPro; IPR03591; IRR_typ.			
DR	Pfam; PF00560; LRR; 14.			
DR	PRINTS; PR00019; LEURICHRPT.			
DR	SMART; SM00082; LRRCT; 1.			
DR	SMART; SM00013; LRRN; 1.			
DR	SMART; SM00369; LRR_TYP; 10.			
KW	Leucine-rich repeat; Repeat; Signal; Transmembrane.			
FT	SIGNAL 1 21	Potential;		
FT	CHAIN 22 581	Leucine-rich repeat-containing protein.		
FT	15.	Extracellular (Potential).		
FT	DOMAIN 22 538	Potential.		
FT	TRANSMEM 539 559	Cytoplasmic (Potential).		
FT	DOMAIN 560 581	LRR 1.		
FT	REPEAT 51 75	LRR 2.		
FT	REPEAT 76 99	LRR 3.		
FT	REPEAT 100 123	LRR 4.		
FT	REPEAT 125 147	LRR 5.		
FT	REPEAT 148 171	LRR 6.		
FT	REPEAT 173 195	LRR 7.		
FT	REPEAT 196 219	LRR 8.		
FT	REPEAT 221 243	LRR 9.		
FT	REPEAT 244 267	LRR 10.		
FT	REPEAT 269 291	LRR 11.		
FT	REPEAT 292 315	LRR 12.		
FT	REPEAT 317 339	LRR 13.		
FT	REPEAT 340 363	LRR 14.		
FT	REPEAT 364 387	LRR 15.		
FT	REPEAT 389 411	N-linked (G1CNAC. . .) (Potential).		
FT	CARBOHYD 75 75	N-linked (G1CNAC. . .) (Potential).		
FT	CARBOHYD 369 369	N-linked (G1CNAC. . .) (Potential).		
SQ	SEQUENCE 581 AA; 64396 MW; 1F3B1485BD2CCB54 CRC64;	SEQUENCE 581 AA; 64396 MW; 1F3B1485BD2CCB54 CRC64;	Query Match	31.3%;
		Best Local Similarity	37.7%;	Score 607;
		Matches 139;	Conservative	Pred. NO. 1.3e-32;
			60;	Mismatches 168;
				Indels 2;
				Gaps 2;
QY	21 CPKTCVKVVRDAACQSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLM 80	21 CPKTCVKVVRDAACQSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLM 80	Query Match	31.3%;
DB	25 CPSECTCRASQVECTGARIYAV-PTPLPNAMSQILNTHITELNESPFLNISALIALR 83	25 CPSECTCRASQVECTGARIYAV-PTPLPNAMSQILNTHITELNESPFLNISALIALR 83	Best Local Similarity	37.7%;
QY	81 LSDSHISAIIDPGTFNDLVKLTKTLRTRNKISRLPRAILDKMVILEQLFLDHNALRDLQDN 140	81 LSDSHISAIIDPGTFNDLVKLTKTLRTRNKISRLPRAILDKMVILEQLFLDHNALRDLQDN 140	Matches 139;	Conservative
DB	84 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	84 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	60;	60;
QY	21 CPKTCVKVVRDAACQSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLM 80	21 CPKTCVKVVRDAACQSGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQRLM 80	Indels	2;
DB	25 CPSECTCRASQVECTGARIYAV-PTPLPNAMSQILNTHITELNESPFLNISALIALR 83	25 CPSECTCRASQVECTGARIYAV-PTPLPNAMSQILNTHITELNESPFLNISALIALR 83	Gaps	2;
QY	81 LSDSHISAIIDPGTFNDLVKLTKTLRTRNKISRLPRAILDKMVILEQLFLDHNALRDLQDN 140	81 LSDSHISAIIDPGTFNDLVKLTKTLRTRNKISRLPRAILDKMVILEQLFLDHNALRDLQDN 140	Indels	2;
DB	84 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	84 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	Gaps	2;
QY	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	Query Match	31.3%;
DB	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	Best Local Similarity	37.7%;
QY	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	Matches 139;	Conservative
DB	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	60;	60;
QY	261 LFLHVSSVSRLLTLPENPLEELPDVLFGEAGLRELLWNLTHLSTLPAAAFFNLSQLTLG 320	261 LFLHVSSVSRLLTLPENPLEELPDVLFGEAGLRELLWNLTHLSTLPAAAFFNLSQLTLG 320	Indels	2;
DB	264 IFMQLPQLNRLLTGFNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 323	264 IFMQLPQLNRLLTGFNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 323	Gaps	2;
QY	141 LFOQLRNLQELGLNQNQISLFLPANLFSSLRLQVLPGLFQQLPQVLR 200	141 LFOQLRNLQELGLNQNQISLFLPANLFSSLRLQVLPGLFQQLPQVLR 200	Query Match	31.3%;
DB	144 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	144 HFSQCSNKLQELQHGNLEYIPDGAFDHVLVGLTNQELALQONQIGLSPGLFQQLPQVLR 203	Best Local Similarity	37.7%;
QY	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	201 LYSNQLTSDGILLSNLGALTELRLERNHILRSVAPGAFDRLGNLSSITLSCGNLLESLPPA 260	Matches 139;	Conservative
DB	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	204 LYENRLTDIPMGTFDGLVNQELNLQKVLQPIGLFQQLPQVLR 203	60;	60;
QY	261 LFLHVSSVSRLLTLPENPLEELPDVLFGEAGLRELLWNLTHLSTLPAAAFFNLSQLTLG 320	261 LFLHVSSVSRLLTLPENPLEELPDVLFGEAGLRELLWNLTHLSTLPAAAFFNLSQLTLG 320	Indels	2;
DB	264 IFMQLPQLNRLLTGFNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 323	264 IFMQLPQLNRLLTGFNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 323	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	Query Match	31.3%;
DB	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	324 LSRN-QISFISPGAFNGLTFLFGNSLKELSIGIFGPMPNRELWLQYDNHISSLPDLNFSNLRQLQVLI 382	Best Local Similarity	37.7%;
QY	381 TLFRNLSSL 389	381 TLFRNLSSL 389	Matches 139;	Conservative
DB	383 NIFANVNGL 391	383 NIFANVNGL 391	60;	60;
QY	383 NIFANVNGL 391	383 NIFANVNGL 391	Indels	2;
DB	RESULT 8	RESULT 8	Gaps	2;
QY	321 LTRNPLRSALPRGVFQGLLRELRVLAELRDDALRGLLQVSLFRENRLRALPR 380	321 LTRN		

OX	NCBI_TaxID=10090;							
RN	[1] SEQUENCE FROM N.A.							
RC	TISSUE=Olfactory epithelium;							
RX	MEDLINE=22388257; PubMed=12477932;							
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubino G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., ToshiYuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulilahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E., Jones S.J., Marra M.A.;							
RA	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";							
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).							
RL	[2] SEQUENCE FROM N.A.							
RN	RP							
RC	TISSUE=Olfactory epithelium;							
RA	Strausberg R.;							
RL	Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.							
DR	EMBL; BC050245; AAH50245.1; -.							
DR	HSSP; Q9BZR6; 1OZN.							
DR	MGD; MGI:1921738; Lrrc15.							
DR	InterPro; IPR01611; IRR.							
DR	InterPro; IPR00483; IRR_Cterm.							
DR	InterPro; IPR00372; IRR_Nterm.							
DR	InterPro; IPR03591; IRR_TYP.							
DR	Pfam; PF00560; LRR; 13.							
DR	PRINTS; PRO00019; LEURICHRPT.							
DR	SMART; SM00082; LRRCT; 1.							
DR	SMART; SM00013; LRRNT; 1.							
DR	SMART; SM00369; LRR_TYP; 11.							
SQ	SEQUENCE 579 AA; 64176 MW; 7841957F6FB643D3 CRC64;							
Query Match	31.0% score 600; DB 2; Length 579;							
Best Local Similarity	36.9%; Pred. No. 3.8e-32;							
Matches	137; Conservative 62; Mismatches 170; Indels 2; Gaps 2;							
Qy	19 FPCPKTCKVVRDAACQCGGGSVAHIAIBGLPNTNLTHILLFRMDQGILRNRHSFSGMTVQLR 78							
Db	23 YGCPSECTCSRASQVECTGAQIVAMPS-PLPWNAMSLQILNTHITELPEDKFLNISALIA 81							
Qy	79 LMLSDSHISAIDEGTFNDLVKLTKLRLTRNKISRLPRAILKMLVILLEQLFDHNALRDL 138							
Db	82 LKMEKELANTIMEGAFRNLSLRLHSLANNKLKNLVPVRLFDVNLTLLSNNQLYQIQ 141							
Qy	139 QNLFQQQLRNQLQELGLNNQNQLSFLPANLSSLRELKLLDSRNNLTHLPKGGLGAQVKLEK 198							
Db	142 PAQFSQFSNLKEIQLQYGNLLEIPEGVFDLVLGTLKLNQNGFTHLSPRVQFHQLGNLQV 201							
Qy	199 LLLYSNQNLTSVDSGLLSNLGALTLERNHLRSVAPGAFDRLGNLSSLTSGNLLESLP 258							
Db	202 LRLYENRRLSDIPMGCTFDALGNLQELAQNLQENQIGTLSPGLFHNRLNQLYLSNNHISHL 261							
Qy	259 PAFFLHVSVSRLTLFENPLEELPVDLGEAGLRELWLNGTHLSTLPLAAFRNLSGLQT 318							
Db	262 PGIFMQLPHLNKLTLFGNSLKELSLSPGVFGMPNLRELWLYNNHITSLSPDNAFSHLNQLQV 321							
Qy	319 LGLTNRNPRLSALPRLGVRQLRVLALHTNALAELRDDALRGGLGHRLROVSLRHNRLFAL 378							
Db	322 LILSHN-QLSYISPGAFNGLTLNRELSLHTNALQDLDGNVFLRNLVSLQNRRLQL 380							
Qy	379 PRTLFRNLSSL 389							
Db	381 PGSIFANVNGL 391							
	RESULT 9							
	LR15_RAT							
	ID LR15_RAT							
	Q8R5M3;							
	AC							
	DT 10-OCT-2003 (Rel. 42, Created)							
	DT 10-OCT-2003 (Rel. 42, Last sequence update)							
	DT 05-JUL-2004 (Rel. 44, Last annotation update)							
	DE Leucine-rich repeat-containing protein 15 precursor (rlib).							
	GN Name=Lrrc15; Synonyms=Lib;							
	OS Rattus norvegicus (Rat).							
	RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.							
	NCBI_TaxID=10116;							
	[1] RN SEQUENCE FROM N.A.							
	RC STRAIN=Wistar; TISSUE=Fetal brain;							
	RC MEDLINE=21645900; PubMed=11785964; DOI=10.1006/bbrc.2001.6272;							
	RA Satoh K., Hata M., Yokota H.;							
	RA "A novel member of the leucine-rich repeat superfamily induced in rat astrocytes by beta-amyloid."							
	RL Biochem. Biophys. Res. Commun. 290:756-762 (2002).							
	CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).							
	CC -1- INDUCTION: By beta-amyloid.							
	CC -1- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.							
	CC							
	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).							
	CC							
	CC DR EMBL; AB071036; BAB84586.1; -.							
	CC DR HSSP; Q9BZR6; 1P8T.							
	CC DR InterPro; IPR001611; LRR.							
	CC DR InterPro; IPR000483; LRR_Cterm.							
	CC DR InterPro; IPR00372; LRR_Nterm.							
	CC DR PFam; PF00560; LRR; 14.							
	CC DR PRINTS; PRO00019; LEURICHRPT.							
	CC DR SMART; SM00013; LRRNT; 1.							
	CC DR SMART; SM00369; LRR_TYP; 11.							
	CC DR Leucine-rich repeat; Repeat; Signal; Transmembrane.							
	CC DR Leucine-rich repeat-containing protein							
	CC DR SMART; SM00082; LRRCT; 1.							
	CC DR SIGNAL 1 21							
	CC FT CHAIN 22 578							
	CC FT DOMAIN 22 535							
	CC FT DOMAIN 536 556							
	CC FT DOMAIN 557 578							
	CC FT REPEAT 51 75							
	CC FT REPEAT 76 99							
	CC FT REPEAT 100 123							
	CC FT REPEAT 125 147							
	CC FT REPEAT 148 171							
	CC FT REPEAT 269 291							
	CC FT REPEAT 292 315							
	CC FT REPEAT 316 339							
	CC FT REPEAT 340 363							
	CC FT REPEAT 364 387							
	CC FT REPEAT 389 411							
	CC FT CARBOHYD 75 75							
	CC FT CARBOHYD 369 369							
	CC SEQUENCE 578 AA; 64127 MW; 991BD057F5912591 CRC64;							

Query Match Score 583; DB 1; Length 578;
 Best Local Similarity 35.9%; Pred. No. 5.3e-31;
 Matches 137; Conservative 62; Mismatches 159; Indels 24; Gaps 4;

Qy 19 FPCPKTCKVVRDAACQCGGSSVHIAELGLPTNL-----THILLFRMDQGILRN 67
 Db 23 YGCPSECTCSRASQVECTGARI-----VAMPTPLPNAMSLQVNTHIT-----ELPE 70

Qy 68 HSFGSGMTVLQRMLSDSHISADPGTFNDLVKLKTLRLTRNKISRLPRAILDKMVLLQL 127
 Db 71 NLFLNISALIALKMEKNEKLSTIMPAGAFRNGLSRLVYSLANNKLRMLPIRVFQDVNLQSL 130

Qy 128 FLDHNAIRDLDQNLFOQLRNQOELGLNQNQNSLFLPANLFLSSLRELKLLDSRNNLTHLPK 187
 Db 131 LLSNNNQLVQIOPAQFSQFNSNRELQHGNNLESIPEEAFAFDHVLGILTKLNLGRNSFTHLSP 190

Qy 188 GLGGAQVKLEKLLYSNQLTSDGSLLSNLGALETBLRERNHLSRSVAPGAFDFRLGNLSSL 247
 Db 191 RLFOQHGNLQVRLHENRLSDIPMGTFDALGNLQELALQENQIGTSLSPGLPHNNRNQLR 250

Qy 248 TLSGNILESLPALFLHVSSSVSRLTLTFENPBLELPDVLFGEAGLREILWLNGLTHLSTLPA 307
 Db 251 YLSNNNHSIQSLPPGIFMQLPQLNKLTFGNSLRELSLSPGVFGPMPNLRWLWYNNHITSLAD 310

Qy 308 AAFRNILSGLQTGLTRNPRLSALPGRGVFGQGLRELRVLAELRDDAELRGLGLHRLQ 367
 Db 311 NTFSHNLNQVLISHN-QLTYISPGAFNGLTLNIRELSLHTNALQDLDNSNVFRSLANLQN 369

Qy 368 VSLRHNRLRALPRTLFRNLLSS 389
 Db 370 ISLQSNRLRQLPGSIFANVNGL 391

RESULT 10

Q9D3K0 PRELIMINARY; PRT; 391 AA.

ID Q9D3K0
 AC Q9D3K0;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
 DE library, clone:5430427N11 product:hypothetical Cysteine-rich flanking
 DE region, C-terminal/Leucine-rich repeat/Leucine-rich repeat, typical
 DE subtype containing protein, full insert sequence.
 Name=Lrcr15;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX Q9D3K0
 RN PRELIMINARY; PRT; 391 AA.

Query Match Score 541.5; DB 2; Length 391;
 Best Local Similarity 35.2%; Pred. No. 2e-28;
 Matches 128; Conservative 59; Mismatches 172; Indels 5; Gaps 3;

Qy 19 FPCPKTCKVVRDAACQCGGSSVHIAELGLPTNLTHILLFRMDQGILRNHSFGMTVLQR 78
 Db 23 YGCPSECTCSRASQVECTGAQIVAMPS-PLPWNAMSQIINTHITELPEDKFLNISALIA 81

Qy 79 LMLDSHISAITDPGTNDLVKLKTLRLTRNKISRLPRAILDKMVLLQLFDHNAIRDLD 138
 Db 82 LKMEKNELANIMPAGAFRNLQSLRHSLSANNKLPVRLFQDVNLQNLVQIQ 141

Qy 139 QNLFOQLRNQELGLNQNQSLFLPANLFLSSLRELKLLDSRNNLTHLPKGLLGAQVKLEK 198
 Db 142 PAQFSQFSNKLQELQLYGNNLEYIPEGVFEDHLVGLTKLNQNLGNNGETHLSPRVFQHGNLQV 201

Qy 199 LLLYSNQLTSDGGLSNNGLGALTELRLELRNHSVAPGAFDRIGNLSSLTLSGNLIESLP 258
 Db 202 LRLYENRLSDIPMGTFDALGNLQELALQENQIGTSLSPGLFHNNRNQLRQLYLSNNHHISHLPLP 261

Qy 259 PALFLHVSSWSRLTLFENPBLELPDVLFGEAGLREILWLNGLTHLSTLPAAAFRNLSGLQT 318
 Db 262 PGFMQOLPHLNKLTLFQNSLKEELSPGVFGPMPNLRELMWLYNNHITSLLPDNAFSLNQLQV 321

Qy	319	LGLTRNPRLSALPRGVFQGLRRELRLVLAELRDDALRGLGHLRQVSLRHNRRLAL	378	DB	27	LLLARLTQPCPVGCDCFGREV-FCSDEQLADIPP-DIPPHITDVFVETAFFTVRTRAF	84
DB	322	LILSHN-QLSYISPGAFANGLTNRELISLHTNALQDDGNNVFRSLANL--GTSHSRITAF	377	Qy	71	SGMTVILQRLMLSDSHISADPGTFNDLVKLTBLTRNKISRLPRAILDMVILLEQLELD	130
Qy	379	PRTL	382	DB	85	SGSPNLTKVFLNTQVRHLEPDAFGGLPRLQDLEITGSPVSNLSAHIFSNLSSLEKLTLD	144
DB	378	DSSL	381	Qy	131	HNALRDLDDQNLFQQLRNQNLQELGLNQNQLSFLPANLESSRELKLLDLSRNRLTHLPKGLL	190
RESULT 11							
Q8R113		PRELIMINARY;	PRT;	562	AA.		
ID	Q8R113;						
AC							
DT	01-JUN-2002	(TREMBLrel.	21,	Created)			
DT	01-JUN-2002	(TREMBLrel.	21,	Last sequence update)			
DT	01-MAR-2004	(TREMBLrel.	26,	Last annotation update)			
DE	1300018K1Rik	protein (Fragment).					
GN							
OS	Mus musculus (Mouse)	.					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;						
NCBI_TAXID	10090;						
[1]							
RN		SEQUENCE FROM N.A.					
RP		STRAIN=FVB/N; TISSUE=Liver;					
RC		MEDLINE=22388257; PubMed=12477932;					
RX		R	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA		R	Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,				
RA		R	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA		R	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA		R	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA		R	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA		R	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA		R	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,				
RA		R	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA		R	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA		R	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA		R	Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,				
RA		R	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA		R	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA		R	Rodriguez A.C., Grimwood J., Schnitz J., Myers R.M., Butterfield Y.S.,				
RA		R	Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E.,				
RA		R	Jones S.J., Marra M.A.; Generation and initial analysis of more than 15,000 full-length human				
RT		RT	and mouse cDNA sequences.;				
RL		RP	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
RN		RP	SEQUENCE FROM N.A.				
RL		RC	STRAIN=FVB/N; TISSUE=Liver;				
RA		RC	Strausberg R.				
RA		RA	Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.				
RA		DR	EMBL; BC025836; AAH25836.1; -.				
RA		DR	HSSP; P07359; 1MOZ.				
RA		DR	MGD; MGI:1919006; 1300018K1Rik.				
RA		DR	InterPro; IPR001611; LRR.				
RA		DR	InterPro; IPR000483; LRR_Cterm.				
RA		DR	InterPro; IPR000372; LRR_Nterm.				
RA		DR	InterPro; IPR003591; LRR_typ.				
RA		DR	Pfam; PF00560; LRR; 12.				
RA		DR	Pfam; PF01462; LRRNT; 1.				
RA		DR	PRINTS; PRO0019; LEURICHRPT.				
RA		DR	SMART; SM00082; LRRCT; 1.				
RA		DR	SMART; SM00013; LRRNT; 1.				
RA		DR	SMART; SM00369; LRR_TYP; 8.				
FT		FT	NON_TER	1	1		
SQ		SQ	SEQUENCE	562	AA;	61993 MW;	C086F0BCFAEF09AE CRC64;
RP		RP	SEQUENCE FROM N.A.				
RC		RC	STRAIN=C57BL/6J; TISSUE=Liver;				
RX		RX	MEDLINE=20499374; PubMed=11042159;				
RA		RA	Carninci P., Shibata Y., Hayatsu N., Suganara Y., Muramatsu M., Hayashizaki Y., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,				
RA		RA	"Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes".				
Query Match	26.0%	Score	503.5;	DB	2;	Length	562;
Best Local Similarity	34.6%	Pred.	No.	1..2e-25;			
Matches 129;	Conservative	69;	Mismatches	172;	Indels	3;	Gaps 3;
Qy	111	LALLRAQQPFPCPKTKCCKVVRDAACCSGGSSVAHIAELGLPTNLTILLERMDOGILRNHSE	70				

RL	Genome Res. 10:1617-1630(2000).	Db	333 TNLSRLVSLTILSHN-ALTDLPEHFRNLEQLWKLSDSNNLTALHPALFHNLSSRLQLLNL 391
RP	SEQUENCE FROM N.A.	QY	371 RHNRLRALPRTLF 383
RC	STRAIN=C57BL/6J; TISSUE=Liver;	DB	392 SRNQLTTLPGGIF 404
RX	MEDLINE=20530913; PubMed=11076861;		
RA	Shibata K., Itoh M., Aizawa K., Nagacka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujikake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuuwa S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.; "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";		
RT	"RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";		
RT	"RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";		
RL	Genome Res. 10:1757-1771(2000).		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Liver;		
RA	Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arakawa T., Bono H., Carninci P., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiracka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito R., Sakai C., Sakaai K., Sano H., Sasak D., Shibata K., Shiba Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.; Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.		
RA	EMBL; AK005049; BAB23775.1; -.		
RA	HSSP; P07359; 1MOZ;		
DR	MGD; MGI:1919006; 1300018K11Rik.		
DR	GO:0004180; F:carboxypeptidase activity; IEA.		
DR	InterPro; IPR01611; IRR.		
DR	InterPro; IPR00483; LRR_Cterm.		
DR	InterPro; IPR00372; LRR_Nterm.		
DR	InterPro; IPR003591; LRR_TYP.		
DR	Pfam; PF00560; LRR; 12.		
DR	Pfam; PF01462; LRRNT; 1.		
DR	PRINTS; PRO00019; LEURICHRPT.		
DR	SMART; SM00082; LRRCT;		
DR	SMART; SM00013; LRRNT; 1.		
DR	SMART; SM00369; LRR_TYP; 10.		
FT	NON_TER 1 1		
SQ	SEQUENCE 557 AA; 61782 MW; A65BFC340CC76563 CRC64;		
	Query Match 23.0%; Score 445; DB 2; Length 557;		
	Best Local Similarity 33.1%; Pred. No. 1e-21;		
	Matches 125; Conservative 54; Mismatches 173; Indels 26; Gaps 4;		
QY	1 MIRSALLSAVLALLRAQPF-PCPKTKCVCVRDAAQCSGGGVAHI-----4.3		
DB	13 MLDGAWLWLWTSLLLARPAQCPGCDCFVQEVF-CSDEELATVPLDIPPYTKNIIIFVET 71		
QY	44 -----AEGLGLPTNLTHILLPRMDQGILRNNHFSQGMVTVLQLMLSDSHISALDPGTEND 96		
DB	72 SFTTLETRAFGSNPNLTKVPLNTQLCQFRPDAFGGLPRLEDLEVTSFLNSTNIFSN 131		
QY	97 LVKLKTLRLTRNKSIRLPRATLDKMLVLEQLFDHNALRDLDQNLFQQLRNLIQELGLNQN 156		
DB	132 LTSLGKLTINFMLEALPEGGLFOHLLQHAALESLHLOGNQLQALPRLFQPLTHLKTNLQAQ 191		
QY	157 QISFLPANLFSLLRELKLLDLSRNLTTHLPKGLLGAQVKEKLLYSNOQLTSVDSGLLSN 216		
DB	192 LIAQLPPEELFPLTSLOTLKLSNNALSGLPQVFGKLGSQELFLDSNNNISELPPQVFSQ 251		
QY	217 LGALTELLRERNHLLRSVAPGAFDRLGNLSSLTLSQELFLDLSNMLSYNAITHLPAQF 276		
DB	252 LFCLERWLLQRNAITHLPLSIFASLGNLTFLSLQWNMLRVLPAGLFATPCLVGLSLLTN 311		
QY	277 PLEELPDVLFGEAGLRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQ 336		
DB	312 QLETVAEGTFAHLSNLMSYNAITHLPAQFIRDLBEEVLYKLYGSN-NLTALHPALFQ 370		
QY	337 GIRELRLVLAHTNAAEL 354		
DB	371 NLSKLELLSLSKNQLTTL 388		
QY	311 RNLSGLQTLGLTRNPRLSALPRGVQGLRELRVLAHTNAAELRDDALRGGLHQLRQVSL 370		

RESULT 14	Q86SU4	PRELIMINARY;	PRT;	560 AA.	DE	Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
ID	Q86SU4;				GN	Name=IGFALS; Synonyms=ALS;
AC					OS	Homo sapiens (Human).
DT	01-JUN-2003	(TREMBLrel.	24;	Created)	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DT	01-JUN-2003	(TREMBLrel.	24;	Last sequence update)	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DT	01-MAR-2004	(TREMBLrel.	26;	Last annotation update)	OX	NCBI_TaxID=9606;
DE	Similar to RIKEN cDNA 1300018K11 gene (Fragment).				RN	[1]
OS	Homo sapiens (Human).				RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
QC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				RC	TISSUE=Liver;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				RX	Medline=92357025; PubMed=1379671;
OX	NCBI_TaxID=9606;				RA	Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RN	[1]				RA	"Structure and functional expression of the acid-labile subunit of the insulin-like growth factor-binding protein complex."
RP	SEQUENCE FROM N.A..				RT	mol. Endocrinol. 6:870-876 (1992).
RC	TISSUE=Colon;				RL	[2]
RA	Strausberg R.;				RN	SEQUENCE FROM N.A..
RL	Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.				RP	Medline=20114388; PubMed=10650966;
DR	EMBL; BC042334; AAH42334.1;	-			RX	Medline=89308584; PubMed=2473065;
DR	HSSP; Q9BZR6; 10ZN.				RA	Baxter R.C., Martin J.L., Beniac V.A. /
DR	InterPro; IPR01611; LRR.				RT	"Conservation of a growth hormone-responsive promoter element in the human and mouse acid-labile subunit genes."
DR	InterPro; IPR00483; LRR_Cterm.				RT	Endocrinology 141:833-838 (2000).
DR	InterPro; IPR00372; LRR_Nterm.				RL	[3]
DR	InterPro; IPR003591; LRR_TYP.				RN	SEQUENCE FROM N.A..
DR	PF00560; LRR; 13.				RP	Frankland J.;
DR	Pfam; PF01463; LRRCT; 1.				RA	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR	Pfam; PF01462; LRRNT; 1.				RL	[4]
DR	PRINTS; PR00019; LEURICHRPT.				RN	SEQUENCE OF 28-35.
DR	SMART; SM00082; LRRCT; 1.				RP	Medline=89308584; PubMed=2473065;
DR	SMART; SM00013; LRRNT; 1.				RX	Medline=264:11843-11848 (1989).
DR	SMART; SM00369; LRR_TYP; 10.				RA	Baxter R.C., Martin J.L., Beniac V.A. /
FT	NON_TER	1			RT	"High molecular weight insulin-like growth factor binding protein complex. Purification and properties of the acid-labile subunit from human serum."
SQ	SEQUENCE 560 AA;	62196 MW;	CABFBBC1E69B0934 CRC64;		RL	J. Biol. Chem. 264:11843-11848 (1989).
Query Match		23.0%	Score 445; DB 2;	Length 560;	CC	- - FUNCTION: Involved in protein-protein interactions that result in protein complexes, receptor-ligand binding or cell adhesion.
Best Local Similarity		33.1%	Pred. No. 1e-21;		CC	- - SUBUNIT: Forms a ternary complex of about 140 to 150 kDa with IGF-I or IGF-II and IGFBP-3.
Matches	125;	Conservative 54;	Mismatches 173;	Indels 26;	CC	- - SUBCELLULAR LOCATION: Extracellular.
QY	1	MLRSALLSAVALLRAQPF-PCPKTKCQVYRDAAQCSGGSVAH1-----		43	CC	- - TISSUE SPECIFICITY: Plasma.
Db	16	MLPGAWLILWTSLLLARPAQCPGCGDCFVQEVF-CSDEELATVPLDIPPYTKNIIFVET 74			CC	- - SIMILARITY: Contains 20 leucine-rich (LRR) repeats.
QY	4	-----AELGLFTNLTHILLFRMDQGILRNHSFSGMTVILQRIMLSDSHISAIDPGTFND 96			CC	-----
Db	75	SFTTLETRAFGSNPNLTKVFLNTQLCQFRPDAFGGLPRLEDTGTSSFLNLSTNIFSN 134			CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions, receptor-ligand binding or cell adhesion.
QY	97	LYKJLKLTRNKNISRLPRAILDKMVLLEQLFDNAILRDLDQNLFOQIIRNLQELQIINQ 156			CC	-----
Db	135	LTSLGKLTNLNFNMLEALPEGGLFOHLLAALESLHQLGNNRLQALPFLRFLQPLTHKTLNLAQN 194			CC	use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)
QY	157	QLSFLPANLFSSSLRELKLLDLSRNNLTHLPKGKLLGAVQVKLEKLLYSNQLTSVDSLGSLSN 216			CC	-----
Db	195	LIAQQLPEELFHPLTSLQRTLKLNSNNALSGLPPQGVFGKLGSQELFLDSNNISELPPQVFSQ 254			DR	EMBL; M86826; AAA36047.1; -.
QY	217	LGALTTELRNHLRSVAPGAFDRLGNLSSLTLSGNLLESLPPALFLHVSSVSRITLFFEN 276			DR	EMBL; AF192554; AAF06774.1; -.
Db	255	LFCLERLWLQRNATHLPLSIFASLGNLTFSLQWNLRLVLPAGLFAHTPCLVGLSLSLN 314			DR	EMBL; AL031724; CAC36078.1; -.
QY	277	PBLELPDVLFGEAGLRELWLNGTHLSTLPAAAFRNLSGLQTIGLTRNPRLSALPRGVFO 336			DR	PIR; A41915; A41915.
Db	315	QLETIVAEGTFAHLSNIRSLMLSNAITHLPAGLFRLDEELVKIYLGSN-NLTALHPALFQ 373			DR	HSSP; P07359; 1M0Z.
QY	337	GLRELRVVLALHTNALAEI 354			DR	Genew; HGNC:5468; IGFALS.
Db	374	NLSKLELLSLSKQQLTTL 391			DR	MIM; 601489; -.
RESULT 15	ALS_HUMAN	STANDARD;	PRT;	605 AA.	DR	GO; GO:0005615; C:extracellular space; TAS.
ID	ALS_HUMAN				FT	GO; GO:0005625; C:soluble fraction; TAS.
AC	P35858;				FT	GO; GO:0005520; F:insulin-like growth factor binding; TAS.
DT	01-JUN-1994	(Rel. 29, Created)			FT	DR; InterPro; IPR001611; LRR_Nterm.
DT	01-JUN-1994	(Rel. 29, Last sequence update)			FT	DR; InterPro; IPR003591; LRR_typ.
DT	05-JUL-2004	(Rel. 44, Last annotation update)			FT	DR; InterPro; IPR000372; LRR_Nterm.
KW	Leucine-rich repeat; Repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Insulin-like growth factor binding protein complex acid labile chain.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Signal.				FT	DR; InterPro; IPR003591; LRR_typ.
KW	Leucine-rich repeat; Glycoprotein;					

FT REPEAT 74 96 LRR 2.
 FT REPEAT 98 120 LRR 3.
 FT REPEAT 121 144 LRR 4.
 FT REPEAT 145 168 LRR 5.
 FT REPEAT 170 192 LRR 6.
 FT REPEAT 193 216 LRR 7.
 FT REPEAT 217 240 LRR 8.
 FT REPEAT 242 264 LRR 9.
 FT REPEAT 265 288 LRR 10.
 FT REPEAT 289 312 LRR 11.
 FT REPEAT 313 336 LRR 12.
 FT REPEAT 337 360 LRR 13.
 FT REPEAT 361 384 LRR 14.
 FT REPEAT 386 408 LRR 15.
 FT REPEAT 409 432 LRR 16.
 FT REPEAT 433 456 LRR 17.
 FT REPEAT 458 480 LRR 18.
 FT REPEAT 482 504 LRR 19.
 FT REPEAT 505 530 LRR 20.
 FT CARBOHYD 64 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 85 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 96 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 368 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 515 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 580 N-linked (GlcNAc. . .) (Potential).
 SQ SEQUENCE 605 AA; 66034 MW; F6562A23CBE918F6 CRC64;

Query Match 21.6%; Score 419.5; DB 1; Length 605;
 Best Local Similarity 33.0%; Pred. No. 5.9e-20;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Qy 38 GSVAH---IAELGLPTNLTHILLPRAILDKMVILLEQQLFIDHNALRDQLNLFQQLRNQELGLN 94
 Db 140 GTBAHTPALASLGLSNN---RLSRLEDGL----FEGLGSLLWDLNLGWNNSLAVLPDAAF 190

Qy 95 NDLYVKLTKLRLTRNKISRLPRAILDKMVILLEQQLFIDHNALRDQLNLFQQLRNQELGLN 154
 Db 191 RGLGSRLRELVLAGNRLAYLQPALFSGLAELRELDLSRNVALRAIKANVFVQLPRLQKLYLD 250

Qy 155 QNQLSFLPANLFSLRELKULLDSRNNNLTQLPK---GLLGAQVKLEKILLYSNQLTSDV 210
 Db 251 RNLIAAVAVGAFLGLKALRWDLSSHNRVAGLLEDTPGGLGLRV---IRLISHNAAIASLR 306

Qy 211 SGLLSNLNGALTELRLERNHILRSVAPGAFDRLGNLSSSLTSGNLLESLPPAFLHVSSVSR 270
 Db 307 PRTFKDHLFEEELQLGHNRIQLAERSFEGLQGLEVLTLDHNQLQEVKAGAFLGLTNVAV 366

Qy 271 LTLFENPLEELPDVLFEGEMAGLRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRN--- 324
 Db 367 MNLSGNCLRNLPQVERGLGKHLHSLLLEGSCLGRPRHTFTGLSGLRLFUKDNGLVGIE 426

Qy 325 -----PRLSALPRGVFQGLRELRLVLAHTNAELRDDAFLRGLGLHILRQ 367
 Db 427 EQSLWGLAELLELDLTSNQLTHLPHRLFQGLGKLEYLLLSRNRLAELPADALGPLQRAFW 486

Qy 368 VSLRHNRRLRALPRTLFRNLSSL 389
 Db 487 LDVSHNRLAEALPNSLLAPLGRRL 508

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:05:23 ; Search time 27 Seconds
 (without alignments)
 955.470 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389

Perfect score: 193.8

Sequence: 1 MLRSALLSAVLALLRAQPFP.....LRHNRRLRALPRTLERNLSSL 389

Scoring table: BLOSUM62

Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : Issued_Patents_AA:
 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:
 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:
 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:
 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:
 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:
 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	76.2	560	3 US-08-592-500-2	Sequence 2, Appli
2	1477	76.2	560	3 US-08-195-006-2	Sequence 2, Appli
3	1477	76.2	560	3 US-09-063-950-4	Sequence 4, Appli
4	1477	76.2	560	5 PCT-US94-07644A-2	Sequence 2, Appli
5	419.5	21.6	605	1 US-08-190-802A-49	Sequence 49, Appli
6	419.5	21.6	605	3 US-08-477-346-49	Sequence 49, Appli
7	419.5	21.6	605	3 US-08-473-089-49	Sequence 49, Appli
8	419.5	21.6	605	4 US-08-487-072A-49	Sequence 49, Appli
9	419.5	21.6	605	4 US-09-538-092-1087	Sequence 1087, Ap
10	416.5	21.5	536	4 US-09-538-092-992	Sequence 992, Ap
11	413.5	21.3	605	3 US-09-063-950-5	Sequence 5, Appli
12	382.5	19.7	603	1 US-08-190-802A-50	Sequence 50, Appli
13	382.5	19.7	603	3 US-08-477-346-50	Sequence 50, Appli
14	382.5	19.7	603	3 US-08-473-089-50	Sequence 50, Appli
15	382.5	19.7	603	4 US-08-487-072A-50	Sequence 50, Appli
16	371.5	19.2	353	3 US-08-986-485-6	Sequence 6, Appli
17	369.5	19.1	1529	4 US-09-312-283C-396	Sequence 396, Appli
18	359.5	18.6	1525	3 US-09-191-647-2	Sequence 2, Appli
19	359.5	18.6	1525	3 US-09-540-245A-2	Sequence 2, Appli
20	359.5	18.6	1525	3 US-09-540-153-2	Sequence 2, Appli
21	357	18.4	1523	3 US-09-182-024A-2	Sequence 2, Appli
22	339	17.5	907	4 US-09-170-496D-264	Sequence 264, Appli
23	339	17.5	907	4 US-09-170-496D-278	Sequence 278, Appli
24	333.5	17.2	1119	4 US-09-907-794A-294	Sequence 294, Appli
25	333.5	17.2	1119	4 US-09-905-125A-294	Sequence 294, Appli
26	333.5	17.2	1119	4 US-09-902-775A-294	Sequence 294, Appli
27	333.5	17.2	1119	4 US-09-906-700-294	Sequence 294, Appli

ALIGNMENTS

28	333.5	17.2	1119	4 US-10-140-002-352	Sequence 352, APP
29	333.5	17.2	1119	4 US-09-903-603A-294	Sequence 294, APP
30	323.5	16.7	1091	3 US-08-986-485-5	Sequence 5, Appli
31	323.5	16.6	642	4 US-10-140-002-370	Sequence 370, App
32	321.5	16.6	620	4 US-09-907-794A-73	Sequence 73, Appli
33	321.5	16.6	620	4 US-09-905-125A-73	Sequence 73, Appli
34	321.5	16.6	620	4 US-09-902-775A-73	Sequence 73, Appli
35	321.5	16.6	620	4 US-09-906-700-73	Sequence 73, Appli
36	321.5	16.6	620	4 US-10-140-002-324	Sequence 324, App
37	321.5	16.6	620	4 US-09-903-603A-73	Sequence 73, Appli
38	319.5	16.5	1480	3 US-09-191-647-7	Sequence 7, Appli
39	319.5	16.5	1480	3 US-09-540-245A-7	Sequence 7, Appli
40	319.5	16.5	1480	3 US-09-540-153-7	Sequence 7, Appli
41	319.5	16.5	1480	3 US-09-182-024A-5	Sequence 5, Appli
42	319.5	16.5	1480	3 PCT-US91-09055-2	Sequence 2, Appli
43	313.5	16.2	708	4 US-09-131-648-2	Sequence 2, Appli
44	313.5	16.2	708	4 US-09-907-794A-69	Sequence 69, Appli
45	313.5	16.2	708	4 US-09-905-125A-69	Sequence 69, Appli

RESULT 1
 US-08-592-500-2
 Sequence 2, Application US/08592500
 Patent No. 6005089
 GENERAL INFORMATION:
 APPLICANT: Lanza, Francois
 APPLICANT: Phillips, David R.
 APPLICANT: Cazenave, Jean-Pierre
 TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Khourie and Crew
 STREET: 379 Lytton Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: US
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/592,500
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/089,455
 FILING DATE: 09-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29,684
 REFERENCE/DOCKET NUMBER: 12418-28
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 326-2422
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 560 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-592-500-2

Query Match 76.2%; Score 1477; DB 3; Length 560;

Best Local Similarity 76.1%; Pred. No. 1.9e-135;
 Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

Qy 1 MIRSALLSAVLALLRAQDPFPCKTCKVVRDAACCSGGSVAHIAELGLPTNLTHILLFRM 60

Db	1	MLRGTLCAVGLLRAQPFPCPACKCVFRDAAQCSGGGDVARISALGLPTNLTHILLFGM	60	Matches	296	Conservative	37	Mismatches	56	Indels	0	Gaps	0
Qy	61	DQGIRLRNHSFSGMTVLQRLMISDSHISAIIDPGTFNDLVRKLTTRNKSIRLPRAILDK	120	Qy	1	MLRSALLSAVILLRAQPFPCPACKCVFRDAAQCSGGGDVARISALGLPTNLTHILLFGM	60	Qy	1	MLRSALLSAVILLRAQPFPCPACKCVFRDAAQCSGGGDVARISALGLPTNLTHILLFGM	60	Qy	1
Db	61	GRGVLTQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLKTRLSRNKITHLPGALLDK	120	Db	1	MLRGTLCAVGLLRAQPFPCPACKCVFRDAAQCSGGGDVARISALGLPTNLTHILLFGM	60	Db	1	MLRGTLCAVGLLRAQPFPCPACKCVFRDAAQCSGGGDVARISALGLPTNLTHILLFGM	60	Db	1
Qy	121	MVLEQFLDHNALRDLDQNLFQQLRNQELGLNQNQLSFLPANLFESSLRELKLLDLSRN	180	Qy	61	DQGILRNHSFSGMTVLQRLMISDSHISAIIDPGTFNDLVRKLTTRNKSIRLPRAILDK	120	Qy	61	DQGILRNHSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLKTRLSRNKITHLPGALLDK	120	Qy	61
Db	121	MVLEQFLDHNALRDLDQNLFQQLRNQELGLNQNQLSFLPASLFTNLLENLKLLDLSGN	180	Db	61	MVLEQFLDHNALRDLDQNLFQQLRNQELGLNQNQLSFLPANLFESSLRELKLLDLSRN	180	Db	61	GRGVLTQSFSGMTVLQRLMISDSHISAVAPGTFSDLIKLKTRLSRNKITHLPGALLDK	120	Db	61
Qy	181	NLTHLPKGLLGQAQVKLEKLLYSNQLTSVDSLRLERNHLSVAPGAFDR	240	Qy	121	MVLEQFLDHNALRDLDQNLFQQLRNQELGLNQNQLSFLPANLFESSLRELKLLDLSRN	180	Qy	121	MVLEQFLDHNALRDLDQNLFQQLRNQELGLNQNQLSFLPASLFTNLLENLKLLDLSGN	180	Qy	121
Db	181	NLTHLPKGLLGQAQAKLERLLHSNRLLVSDGILNLNQLDFLPASLFTNLLENLKLLDLSGN	180	Db	121	NLTHLPKGLLGQAQVKLEKLLYSNQLTSVDSLRLERNHLSVAPGAFDR	240	Db	121	NLTHLPKGLLGQAQAKLERLLHSNRLLVSDGILNLNQLDFLPASLFTNLLENLKLLDLSGN	180	Db	121
Qy	241	LGNLSSLTLSGNLLESPPALEFLHVSSVSRLTIFENPLEELPDVLFGEMAGLRELWLNGT	300	Qy	181	NLTHLPKGLLGQAQVKLEKLLYSNQLTSVDSLRLERNHLSVAPGAFDR	240	Qy	181	NLTHLPKGLLGQAQVKLEKLLYSNQLTSVDSLRLERNHLSVAPGAFDR	240	Qy	181
Db	241	LPNLSSLTLSRNHLAFLPSALFLSHNLITLTLIFENPLAELPGVLFGEMGGQELWLNR	300	Db	181	NLTHLPKGLLGQAQAKLERLLHSNRLLVSDGILNLNQLDFLPASLFTNLLENLKLLDLSGN	180	Db	181	NLTHLPKGLLGQAQAKLERLLHSNRLLVSDGILNLNQLDFLPASLFTNLLENLKLLDLSGN	180	Db	181
Qy	301	HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFOGLRELRLVLAELRDDALR	360	Qy	241	LGNLSSLTLSGNLLESPPALEFLHVSSVSRLTIFENPLEELPDVLFGEMAGLRELWLNGT	300	Qy	241	LGNLSSLTLSGNLLESPPALEFLHVSSVSRLTIFENPLEELPDVLFGEMAGLRELWLNGT	300	Qy	241
Db	301	QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFQGLGELQVLAHSNGLTALPDGLLR	360	Db	241	LPNLSSLTLSRNHLAFLPSALFLSHNLITLTLIFENPLAELPGVLFGEMGGQELWLNR	300	Db	241	LPNLSSLTLSRNHLAFLPSALFLSHNLITLTLIFENPLAELPGVLFGEMGGQELWLNR	300	Db	241
Qy	361	GLGHLRQVSLRHNRLRALPRTLFRNLSSL	389	Qy	301	HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFOGLRELRLVLAELRDDALR	360	Qy	301	HLSTLPAAAFRNLSGLQTGLTRNPRLSALPRGVFOGLRELRLVLAELRDDALR	360	Qy	301
Db	361	GLGKLRLRQVSLRHNRLRALPRTLFRNLSSL	389	Db	301	QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFQGLGELQVLAHSNGLTALPDGLLR	360	Db	301	QLRTLPAAAFRNLSRLRYLGVTLSPRLSALPQGAFQGLGELQVLAHSNGLTALPDGLLR	360	Db	301

RESULT 2
US-08-195-006-2

RESULT 2
US-08-195-006-2

Sequence 2, Application US/08195006
ent No. 6083688
VERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTY: San Mateo

COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #11.0, V
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: DOW, Karen B.
REGISTRATION NUMBER: 29,684
PREFERENCE/DOCUMENT NUMBER: 12418-28

REFERENCE/DOCID: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
FORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 560 amino acid
TYPE: amino acid
TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-195-006-2

94

RESULT 3
US-09-063-950-4
; Sequence 4, Application US/09063950C
; Patent No. 6225085
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFORE
; FILE REFERENCE: MEI-019
; CURRENT APPLICATION NUMBER: US/09/063,950C
; CURRENT FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-063-950-4

Query Match	Score	DB	Length
Best Local Similarity	76.18	Pred. No.	1.9e-135;
Matches 296; Conservative	37; Mismatches 56; Indels 0; Gaps 0;		
Qy			
1	MLRSALLSAVLALLRAQQFPCKVVRDAACSGGSSVAHIAELGLPFTNLTHILLFRM	60	
1	MLRGTLCAVGLLRAQQFPCKVFRDAACSGGDDVARISALGLPFTNLTHILLFGM	60	
Qy			
61	DQGILRNHSFGMTVLRQLMLSDSHISAJDPGTFNLDLVKLKTLRLLTRNKISRLPRAILDK	120	
61	GRGVILQSOFGMTVLRQLMISDSHISAVAPGTFSDLIKLKTLRLLSRNKJITHLPGALLDK	120	
Qy			
121	MVLLEQFLDHNALRDQNLFQQLRLNQELGLNNQNQLSFLPANLFFSSLRELKLLDLSRN	180	
121	MVLLEQFLDHNALRGIDQNMFQKLVNLQELALNNQQLDFLPASLFTNLENLKLDDLSGN	180	
Qy			
181	NLTHLPKGGLLGAQVKLEKLLLNSNQLTSVDSGLLSNLGALETLLERNHILRSVAPGAFDR	240	
181	NLTHLPKGGLLGAQAKLERLLHSNRLVSLDSGLLNSLGALETQFHRNHIRSIAPGAFDR	240	
Qy			
241	LGNLSSLTISGNLLESIIPPALEFLHVSSVSRLLTLEENPLEELPDVLFGEAGLRELWLNGT	300	
Qy			

301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPGRGVFQGLRELRLVLAHTNALAELRDDALR 360
 301 QLRTLPAAAFRNLSRLRYLGVTLSPLRSALPQGAFQGLGELQVLALHSNGLTALPDGLLR 360
 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSL 389
 361 GLGKLRQVSLRRNRRLRALPRTLFRNLSL 389

RESULT 5
 US-08-190-802A-49
 Sequence 49, Application US/08190802A
 Patent No. 5519003
 GENERAL INFORMATION:
 APPLICANT: COR Therapeutics, Inc.
 TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Khourie and Crew
 STREET: 379 Lytton Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: US
 ZIP: 94301
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/07644A
 CLASSIFICATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29, 684
 REFERENCE/DOCKET NUMBER: 012418-003000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 326-2400
 TELEFAX: (415) 326-2422
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 560 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US94-07644A-2

Query Match 76.2%; Score 1477; DB 5; Length 560;
 Best Local Similarity 76.1%; Pred. No. 1.9e-135;
 Matches 296; Conservative 37; Mismatches 56; Indels 0; Gaps 0;

QY 1 MLRSELLSAVLLRAQPFPCPKTCKCVRDAAOQCGGGSVAHIAELGLPTNLTHILLFRM 60
 DB- 1 MLRGTLCAVLGILRAQPFPCPACKCVFRDAAQCGGGDVARIISALGLPTNLTHILLFGM 60

QY 61 DQGILRNHSFSGMTVLQRLMLSDSHISAIIDPGTFNDLVKLTTRLTRNKSISRLPRAILDK 120
 DB- 61 GRGVLQSQFSGMTVLQRLMISDSHISAVAPGTFSDLKLTKTSLRNKITHLPGALLDK 120

QY 121 MVLLEQLFLDHNALRDLQNLFOQLRNQELGLNQNQNLQELALNQNQLDLFLPASLFTNLLENLKLDLSGN 180
 DB- 121 MVLLEQLFLDHNALRGIDQMFQKLVLNQELALNQNQLDLFLPASLFTNLLENLKLDLSGN 180

QY 181 NLTHLPKGLLGAQVKLEKLLYSSNQLTSDVSGLSSNLGAALTTELRLERNHRSVAPGAFDR 240
 DB- 181 NLTHLPKGLLGAQAKLERLLHSNRLVSDGLLNSLGAALTTELQFHRNHRSVAPGAFDR 240

QY 241 LGNLSLTLGGNLLSLLPAPFLHVSSVSRLLTENPLFVQFLGEMGQELWLNRT 300
 DB- 241 LPNLSSLLTSRNRHAFLPSALFLHSHNLTLTFENPLFVQFLGEMGQELWLNRT 300

QY 301 HLSTLPAAAFRNLSGLQTGLTRNPRLSALPGRGVFQGLRELRLVLAHTNALAELRDDALR 360

Db 301 QLRTLPAAAFRNLSRLRYLGVTLSPLRSALPQGAFQGLGELQVLALHSNGLTALPDGLLR 360
 Qy 361 GLGHLRQVSLRHNRRLRALPRTLFRNLSL 389
 Db 361 GLGKLRQVSLRRNRRLRALPRTLFRNLSL 389

RESULT 5
 US-08-190-802A-49
 Sequence 49, Application US/08190802A
 Patent No. 5519003
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: P.O. Box 60850
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-0850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/190,802A
 FILING DATE: 01-FEB-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 605 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Insulin-like growth factor binding protein complex, Fig. 32
 US-08-190-802A-49

Query Match 21.6%; Score 419.5; DB 1; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Qy 38 GSVAH---IAEGLGLPTNLTHILLFRMDQGILRNHSFSGMVTVLQRLMLSDSHISAIIDPGTF 94
 Db 140 GTFAAHTPALASLGSNN---RISRLEDGL---FEGLSLWDLNIGWNSLAVLPDAAF 190

Qy 95 NDLVVKLTKTLLTRNKSISRLPRAILDRDLQNLFOQLRNQELGLN 154
 Db 191 RGLGSLRELVLAGNRILAYLQPAFSGLAELRELDLSRNALRAIKANVFVQLPRLQKLYLD 250

Qy 155 QNQLSFLPANLFSSLREKLKUDLDSRN 180
 Db 211 SGLLSNLGALTELRLERNHRSVAGLLEDTFPGLLGLRV---LRISHNIAASLR 306

Qy 211 RNLLAAVAPGAFGLKALRWDLISHNRVAGLLEDTFPGLLGLRV---LRISHNIAASLR 306

Db 307 PRTFKDLHFEELQGLHNRIRQLAERSFEGLGQLEVLTIDHNQLOQEVKAGAFLGLTNVAV 366

QY 271 LTLFENPLEELPDVLFGEAGLRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRN----- 324
 Db 367 MNLSGNCLRNLPSEQVFRGLGKHLHSIHLLEGSCLGRIRPHTFTGSLQRLKDNGLVGE 426
 QY 325 -----PLSALPRGVFQGLRELRVLALHTNALAELRDDALRGLGHLRQ 367
 Db 427 EQSLWGLAELLELDLTSNQLTHLPHRLFQGLGKLEYLLSSRNRLAELPADALGPLQRAFW 486
 QY 368 VSLRHNRRLRALPRTLFRNLSL 389
 Db 487 LDVSHNRLEALPNSSLAPLGR 508

RESULT 6
 US-08-477-346-49
 Sequence 49, Application US/08477346
 Patent No. 6262023

GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 605 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 INDIVIDUAL ISOLATE: Insulin-like growth factor binding
 INDIVIDUAL ISOLATE: protein complex, Fig. 32
 US-08-477-346-49

Query Match 21.6%; Score 419.5; DB 3; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

QY 38 GSVAH---IAEIGLPLPTNLTHLLFRMDQGILRNHSFSGMTVILQRLMLSDSHISAIIDPGTF 94
 Db 140 GTFAHTPALASLGLSNN---RLSRLDEDGL---FEGLGLSLWDLNLGWNNSLAVLPDAAF 190
 QY 95 NDLYKLTKLRLTRNKISRLPRAILDRMVLLEQFLDHNALRDQNLFOQLRNQELGIN 154
 Db 191 RGLGSLRELVLAGNRLAYLQPAFSGLAELRELDLSRNALRAIKANTVFOQLPRLQKLYLD 250

RESULT 7
 US-08-473-089-49
 Sequence 49, Application US/08473089
 Patent No. 6342368

GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-1500
 INFORMATION FOR SEQ ID NO: 49:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 605 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 INDIVIDUAL ISOLATE: Insulin-like growth factor binding
 INDIVIDUAL ISOLATE: protein complex, Fig. 32
 US-08-473-089-49

Query Match 21.6%; Score 419.5; DB 3; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

QY 38 GSVAH---IAEIGLPLPTNLTHLLFRMDQGILRNHSFSGMTVILQRLMLSDSHISAIIDPGTF 94
 Db 140 GTFAHTPALASLGLSNN---RLSRLDEDGL---FEGLGLSLWDLNLGWNNSLAVLPDAAF 190
 QY 95 NDLYKLTKLRLTRNKISRLPRAILDRMVLLEQFLDHNALRDQNLFOQLRNQELGIN 154
 Db 191 RGLGSLRELVLAGNRLAYLQPAFSGLAELRELDLSRNALRAIKANTVFOQLPRLQKLYLD 250

Query Match 21.6%; Score 419.5; DB 3; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

QY 38 GSVAH---IAEIGLPLPTNLTHLLFRMDQGILRNHSFSGMTVILQRLMLSDSHISAIIDPGTF 94

Db 140 GTFAHTPALASLGNSN---RLSRLEDGL----FEGLGSWDLNLGWNSLAVLPDAAF 190
 QY 95 NDLVKKTLLTRNKISRLPRAILDKMVILLEQFLDHNALRDLDQNLFOQLRNLOEGLLN 154
 Db 191 RGLGSRELVLAGNRAYLQPALFSGLAAELRELDLSRNALRAIKANVFVQLPRLQKLYD 250
 QY 155 QNQLSFLPANLPESSLREKLDDLSRNNLTHLPK---GLLGAQVKLEKLLYSNQLTSDV 210
 Db 251 RNLIAVAAPGAFLGLKALRWLDLSHNRVAGLLEDTEPGILLGLRVR---LRLSHNAIASLR 306
 QY 211 SGLLSNLGALTBLRERNHLSRSVAPGAFDRIGNLSSLTSGNLLSLLPALLFLHVSSVSR 270
 Db 307 PRTFKDLHFILELLQJGHNRIRQLAERSFEGIGQLEVITLDDHNLQEVKAAGAFLGTVNAV 366
 QY 271 LTLFENPLEELPDVLFEGMAGLRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRN----- 324
 Db 367 MNLSGNCLRNLPQEVRGLGKLHSUHLEGSCLGRIRPHTFTGIGSLRRFLKNDNGLVGIE 426
 QY 325 -----PRLSALPRGVFOGLRRELRLVLAELRDDALRGLGLHLRQ 367
 Db 427 EQSLWGLAELLELDTSNQLTTHLPHRLFOGLKLEYLLSSRNRLAELPADALGPLQRAFW 486
 QY 368 VSLRHNRRLRALPRTLFRNLSSL 389
 Db 487 LDVSHNRLEALPNSLLAPLGR 508

RESULT 8
 US-08-487-072A-49
 ; Sequence 49, Application US/08487072A
 ; Patent No. 6423684
 ; GENERAL INFORMATION:
 ; APPLICANT: Ron, Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morrison & Foerster
 ; STREET: 2000 Pennsylvania Avenue, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/487,072A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29, 959
 ; REFERENCE/DOCKET NUMBER: 2550-0025.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; INFORMATION FOR SEQ ID NO: 49:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 605 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: Insulin-like growth factor binding protein complex, Fig. 32
 ; INDIVIDUAL ISOLATE: protein complex, Fig. 32
 ; US-08-487-072A-49

Query Match 21.6%; Score 419.5; DB 4; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Db 38 GSVAH---IAELGLPPTNLTHILLFRMDQGILRNHSFSGMTVLQRLMLSDSHISADPGTF 94
 QY 140 GTFAHTPALASLGNSN---RLSRLEDGL----FEGLGSWDLNLGWNSLAVLPDAAF 190
 Db 95 NDLVKKTLLTRNKISRLPRAILDKMVILLEQFLDHNALRDLDQNLFOQLRNLOEGLLN 154
 QY 191 RGLGSRELVLAGNRAYLQPALFSGLAAELRELDLSRNALRAIKANVFVQLPRLQKLYD 250
 Db 155 QNQLSFLPANLPESSLREKLDDLSRNNLTHLPK---GLLGAQVKLEKLLYSNQLTSDV 210
 QY 95 NDLVKKTLLTRNKISRLPRAILDKMVILLEQFLDHNALRDLDQNLFOQLRNLOEGLLN 154
 Db 191 RGLGSRELVLAGNRAYLQPALFSGLAAELRELDLSRNALRAIKANVFVQLPRLQKLYD 250
 QY 155 QNQLSFLPANLPESSLREKLDDLSRNNLTHLPK---GLLGAQVKLEKLLYSNQLTSDV 210
 Db 251 RNLIAVAAPGAFLGLKALRWLDLSHNRVAGLLEDTEPGILLGLRVR---LRLSHNAIASLR 306
 QY 211 SGLLSNLGALTBLRERNHLSRSVAPGAFDRIGNLSSLTSGNLLSLLPALLFLHVSSVSR 270
 Db 307 PRTFKDLHFILELLQJGHNRIRQLAERSFEGIGQLEVITLDDHNLQEVKAAGAFLGTVNAV 366
 QY 271 LTLFENPLEELPDVLFEGMAGLRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRN----- 324
 Db 367 MNLSGNCLRNLPQEVRGLGKLHSUHLEGSCLGRIRPHTFTGIGSLRRFLKNDNGLVGIE 426
 QY 325 -----PRLSALPRGVFOGLRRELRLVLAELRDDALRGLGLHLRQ 367
 Db 427 EQSLWGLAELLELDTSNQLTTHLPHRLFOGLKLEYLLSSRNRLAELPADALGPLQRAFW 486
 QY 368 VSLRHNRRLRALPRTLFRNLSSL 389
 Db 487 LDVSHNRLEALPNSLLAPLGR 508

RESULT 9
 US-09-538-092-1087
 ; Sequence 1087, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9
 ; SEQ ID NO: 1087
 ; LENGTH: 605
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0) .. (0)
 ; OTHER INFORMATION: Polypeptide Accession Number P35858
 ; US-09-538-092-1087

Query Match 21.6%; Score 419.5; DB 4; Length 605;
 Best Local Similarity 33.0%; Pred. No. 3.7e-32;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Db 38 GSVAH---IAELGLPPTNLTHILLFRMDQGILRNHSFSGMTVLQRLMLSDSHISADPGTF 94
 QY 140 GTFAHTPALASLGNSN---RLSRLEDGL----FEGLGSWDLNLGWNSLAVLPDAAF 190
 Db 95 NDLVKKTLLTRNKISRLPRAILDKMVILLEQFLDHNALRDLDQNLFOQLRNLOEGLLN 154
 QY 191 RGLGSRELVLAGNRAYLQPALFSGLAAELRELDLSRNALRAIKANVFVQLPRLQKLYD 250

Qy 155 QNQLSFPLANLFSSSLREKLKLLDSLRSNNLTHLPK---GLLGAQVKLEKULLYSNQLTSVWD 210
 Db 251 RNLIAAVAPGAFLGLKALRWLDSLHNRVAGLLEDTFPGLLGLRV---LRLSHNAIASLR 306

Qy 211 SGLLSNLGALTELLERNLHRSVAPGAFLDRGNLSSLTLSGNLLQESLPPALFLHVSSVSR 270
 Db 307 PRTFEDLHFLFELQIGHNTRIOLAERSFEGIQLQEVLTLDHNLQEVKAGAFLGLTNNAV 366

Qy 271 LTLFENPPEELPDVLFGEAGLRLREWLNGTHLSTLPAAFRNLSGLQTLGLTRN--- 324
 Db 367 MNLSGNCLRNLPSEQVFRGLGKHLHLEGSCLGIRPHTFTGLSGLRRRLFLKDNGLVIE 426

Qy 325 -----PRLSALPGRGVFQGLLRELRVLAHTNNAELRDDALRGLGLHLRQ 367
 Db 427 EQSLWGLAELLELDLTSNQNLTHLPHRLFQGLGKLEYLLSRRNRLAELPADALGPLQRAFW 486

Qy 368 VSLRHNRNRLRALPRTLFERNLSSL 389
 Db 487 LDVSHNRNLEALPNSLLAPLGRLL 508

RESULT 10
 US-09-538-092-992
 ; Sequence 992, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; INVENTION: Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538, 092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127, 352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178, 965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuraPaseqFormatter Version 0.9
 ; SEQ ID NO: 992
 ; LENGTH: 536
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0) (0)
 ; OTHER INFORMATION: Polypeptide Accession Number P22792
 US-09-538-092-992

Query Match 21.5%; Score 416.5; DB 4; Length 536;
 Best Local Similarity 32.3%; Pred. No. 6.1e-32;
 Matches 115; Conservative 53; Mismatches 163; Indels 25; Gaps 3;

Qy 22 PRTCKCVRDAQCGSGSVAH-----AEIQLPTNLTHLIF 58
 2 PMGDCDFVQEVF-CDSEELATVPLDIPPYTKNLIIFVETSFTTLETRAFGSNPNLTKVVF 60

Qy 59 RMDQGILRNHSFSGMTVQLRIMLSDSHISAIDPGTFLNDLVKLKTLRTRNKISRLPRAIL 118
 61 DTQLCQFRPDAFGGLPRLEDIETVGSSFLNLSTNIFSNLTSGLKLTLNFMNMLEALPEGLF 120

Qy 119 DKMVLLEQLFDHNAIRLDQNLFQQLRNQESLFLPANLFSSRELKLLDSL 178
 Db 121 QHLAALESLHLOGNQLQALPRLFQPLTHLKTLNIAQNLLAQPLPEELFHPPLTSQTLKLS 180

Qy 179 RNNLTHLPKGLLGQAQVKLEKULLYSNQLTSVDSLGSLLNLGALTBLRERNHRSVAPGAF 238
 Db 181 NNALSGCLPQGVFQKGLQELFLDSNNNISELPQVFSQLFCLERIWLQRNAITHPLSIF 240

Qy 239 DRIGNLSSSLTSGNLLLESLPPALFLHVSSVSRLTFLFENPPEELPDVLFGEMAGLRELWLN 298
 Db 241 ASIGNLTFLSSLQWNMRLVPLPAGLFAHTPCLVGLSLTHNOLETVEGTFAHLSNRLSMLMS 300

Qy 299 GTHLSTLPAAAFRNLNSGLQTLGLTRNPRLSALPRGVFQGLRELRVLAHTNNAEL 354

Db 301 YNAITHLPAGIFRDLLEELVVKLYLGSN-NLTALHPALFQNLSKLELLSLSKQNLTTL 355

RESULT 11
 US-09-063-950-5
 ; Sequence 5, Application US/09063950C
 ; Patent No. 6225085
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; INVENTION: NOVEL LRG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFORE
 ; FILE REFERENCE: MBI-019
 ; CURRENT APPLICATION NUMBER: US/09/063, 950C
 ; CURRENT FILING DATE: 1998-04-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO: 5
 ; LENGTH: 605
 ; TYPE: PRT
 ; ORGANISM: Papio hamadryas
 US-09-063-950-5

Query Match 21.3%; Score 413.5; DB 3; Length 605;
 Best Local Similarity 29.8%; Pred. No. 1.4e-31;
 Matches 141; Conservative 56; Mismatches 167; Indels 109; Gaps 8;

Qy 21 CPKTKCVCVRDAAQ-----CSG-----
 41 CPATCACSYDDEVNELSVFCSSRNLTLPDPGIPPGTQALWLDSDSHIS 100

Db 38 -----GSVAHIAELGILPNTNLTHILLFRMDQGILRNHSFSGMTVILQRLMLSDSHIS 87
 101 AFLNLQGGQQLGSLEPQALLGL-ENLCHLHERNQLRSLAVGTFAYTPALLGNSNNRLS 159

Qy 88 AIDPGTFNDLVKLKTLRLTRNKISRLPRAI-----LDMVVL-----
 160 RIEDGLFEGLGNLWDLNQGWNSLAVLPDAAFRGGLGRELVLAVGRILAYLQPALFSGLAE 219

Db 124 LEQLFLDHNALRDQNLFOOLRNQELGLNQNQNSFLPANLFSSELRLKLLDSLRSRNLT 183
 220 LRELDLSRNALRAIKANVFAQLPRLQKLYLDRNLIAAVAPGAFLGLKALRWLDLSHNRYA 279

Qy 184 HLPK-----GILGAQVKLEKLLYSNQLTSVDSGLLSNLGALTBLRERNHRSVAPGAFD 239
 Db 280 GILLEDTFPGLLGLRV---LRLSHNAIASLRRPTFEDLHFLEELQLGHNRIQLAERSFE 335

Qy 240 RICGNLSSLTSGNLLESLPPALFLHVSSVSRLTFLFENPPEELPDVLFGEMAGLRELWLN 299
 Db 336 GIGGQLEVLTLDHNQLQEVKVGAFLGLTNVAVMNLSGNCLRNLPQVFRGLGKLNLSHLEG 395

Qy 300 THLSTLPAAAFRNLSGQLTGLTRN-----
 396 SCLGRIRPHTFAGLSGLRRLFLKDNGLYGIEQSLWGLAELLELDLTSNQLTHLPHQLFQ 455

Db 337 GLRELRLVLAHTNNAELRDDALRGLGHLRQVSLLRNRRLALPRTLFRNLSSL 389
 Db 456 GLGKLEYLLSHNRLAELPADALGPIQRAFWMDVSHNRLAELPGSLLASLGRL 508

RESULT 12
 US-08-190-802A-50
 ; Sequence 50, Application US/08190802A
 ; Patent No. 5519003
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron, Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: P.O. Box 60850

CITY: Palo Alto
 STATE: CA
 ZIP: 94306-0850
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/190,802A
 FILING DATE: 01-FEB-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 603 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
 INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
 US-08-190-802A-50

Query Match 19.7%; Score 382.5; DB 1; Length 603;
 Best Local Similarity 28.7%; Pred. No. 1.5e-28;
 Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

Qy 21 CPKTCKCVRD----AAQCGGSVAHIAELGLPTNLTHILLFRMDQGILRNHSFGMTV 75
 Db 41 CPVACTCSHDDYTDELSVFCSKKNLTHLPD-DIPVSTRALWLDGNNLSSIPSAAFQNLSS 99

Qy 76 LQRLMLSDSHISAIDPGTFNDLVKLKTRLTRNKRSLRPAILDKMVILLEQLFDHNALR 135
 Db 100 LDFLNQGSWLRSLPQALLGLQNLYYHLERNLRNLAVGLFTHTPSLASLSSNLIG 159

Qy 136 DLDQNLFQQLRNQOELGLNNQNLSEFLPANLFFSSLRLKLLDLSRNNLTHLPKGLLGAQVK 195
 Db 160 RLEEGLFQQLSHLWDLNGLWNSLVLVLPDTVFQGLGNLHEVLAGNKLTYLQPAFLCGIGE 219

Qy 196 LEKLLLYSNQLTSDVDSGLLNLGALTELRNHLRSVAPGAF-----DRLG 242
 Db 220 LRELDLSRNALRSVKANVVFHLPRLQKLYLDRNLITAVAPGAFGLGMKALRWLDLSHNRVA 279

Qy 243 NLSS-----LTLISGNLLESLPPALFLHVS SVSRLTFLFENPLEELPDVLFEGMAG 291
 Db 280 GLMEDTFPGLLGHLVRLAHNAIASLRPRTFKDLHFLFELQGLHNRIRQLGERTFEGLQ 339

Qy 292 LREIWLNGTHLSTLPPAAAFLRNGLQTLGLTRNPRLSALPRGVFOGLRELRLV-----344
 Db 340 LEVLTDNDQITEVRVGAFLQSLAGLSEELLDLTNRNRLTHLPQLFQGLG 398

Qy 345 ---ALHT-----NALAELRDDALRGLQHRLRQVSRLRNRLRALPRTLFRNLS 387
 Db 399 GHVRLHTFAGLSGLRRLFLRDNSISSIEQSLAGLSEELLDLTNRNRLTHLPQLFQGLG 458

Qy 388 SL 389
 Db 459 HL 460

RESULT 13
 US-08-477-346-50
 ; Sequence 50, Application US/08477346

Qy 292 LREWLNGTHLSTLPPAAAFLRNGLQTLGLTRNPRLSALPRGVFOGLRELRLV-----344
 Db 340 LEVLTDNDQITEVRVGAFLQSLAGLSEELLDLTNRNRLTHLPQLFQGLG 398

345 ---ALHT----NAAELRDDALRGLGHLRQVSRLRNRLALPRTLFRNLS 387
 399 GHVRLHTFAGLSSGRLRFLRDNISSEEQSLAGLSEELLDLTTNRLTHLPRQLFGLG 458
 388 SL 389
 459 HL 460

Sequence 50, Application US/08473089
 Patent No. 6342368

GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 603 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
 INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
 S-08-473-089-50

RESULT 14
 -08-473-089-50
 Sequence 50, Application US/08473089
 Patent No. 6342368

GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,072A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 603 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: Insulin-like growth factor bind.
 INDIVIDUAL ISOLATE: pro. complex-rat, Fig. 33
 S-08-473-089-50

Query Match 19.7%; Score 382.5; DB 3; Length 603;
 Best Local Similarity 28.7%; Pred. No. 1.5e-28;
 Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

21 CPKTCKCVVRD----AAQCGGSSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTV 75
 41 CPVACTSHDDYTDELSVFCSSKNLTHLP-DIPVSTRALWLDGNLNLSSIPSAAFQNLS 99

76 LQRLMLSDSHISAIDPGTFNDLVKLKTLRTRNKISRLPRAILDKMVLQLFDHNAIR 135
 100 LDFLNLOGSWLRSLEPQALLGLQNLYYLHERRNRLRNLAvgLFTHTPSLASLSSNLLG 159

136 DLQDNLFLQQQLNLQELGLQNQNLSPANLSSLRKLDDLSRNRLTHLPKGLLGAQVK 195
 160 RLEEGLFQGGLSHLDNLGWNSLUVLPDTVFGQLGNLHELVAGNKLTYLQPALEFCGLGE 219

196 LEKLLLYSNQLTSDGLLSNLGALETLLERNHLSVAPGAF-----DRLG 242

Query Match 19.7%; Score 382.5; DB 4; Length 603;
 Best Local Similarity 28.7%; Pred. No. 1.5e-28;
 Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

21 CPKTCKCVVRD----AAQCGGSSVAHIAELGLPTNLTHILLFRMDQGILRNHSFSGMTV 75
 41 CPVACTSHDDYTDELSVFCSSKNLTHLP-DIPVSTRALWLDGNLNLSSIPSAAFQNLS 99

Y b 100 LDFLNLOGSWLRSLEPQALLGLQNLYYLHERRNRLRNLAvgLFTHTPSLASLSSNLLG 159

Y b 136 DLQDNLFLQQQLNLQELGLQNQNLSPANLSSLRKLDDLSRNRLTHLPKGLLGAQVK 195

Y b 160 RLEEGLFQGGLSHLDNLGWNSLUVLPDTVFGQLGNLHELVAGNKLTYLQPALEFCGLGE 219

Y b 196 LEKLLLYSNQLTSDGLLSNLGALETLLERNHLSVAPGAF-----DRLG 242

QY 243 NLSS----LTLSGNLLESPPALFLFHSSVSRLTLEPVLFGEMAG 291
 Db 280 GLMEDTPGPGLLGHVRLAHNIAASLRRPTFKDLHLEELQLGHNRIRQLGERTFEGLGQ 339

QY 292 LRELWLNGLHISTLPAAAFRNLSLQLTGLGLTRNPRLSALPRGVFQGLRELRVL---- 344
 Db 340 LEVLTLDNQITEVRVGAFAFSGLFVNVAVMNLSGN-CLRSLPERVFGQLDKLHLSLHLSCL 398

QY 345 ---ALHT----NALAELRDDALRGLGLHLRQVSLRHNRLRALPRTLFRNLS 387
 Db 399 GHVRLHTFAGLSSGRLRFLRDNISSEEQSLAGLSEELLDLTTNRLTHLPRQLFQGLG 458

QY 76 LQRLMLSDSHISAIDPGTFNDLVKLTKTLRTRNKISRUPRAILDKMVLLEQLFLDHNAIR 135
 Db 100 LDFLNQGSWLRSLPQALLGLQNLYYLHLRNRLRNLAUGLFTHTPSLASLSLSSNNLG 159
 QY 136 DLDQNLFQQQLRNLQELGLNQNQNLSPFLPANLFSLRLKLLDLSRNNNLTHLPKGILLGAQVK 195
 Db 160 RLEEGLFQGLSHLWDLNLGWNSLVLVLPDTVQGLGNLHLVLAGNKLTYLQPALFCGLGE 219
 QY 196 LEKLLYSNQLTSDSGLLNLGALTELRERNHRSVAPGAF-----DRLG 242
 Db 220 LRELDLSRNALRSVKANVFVHLPRLQKLYLDRLNLTAVAPGAFGLGMKALRWLDLSHNRVA 279
 QY 243 NLSS-----LTLSGNLLESLPALFLHVSSVSRLLTFLFENPLEELPDVLFGEVAG 291
 Db 280 GLMEDTFPGLLGLHVLRLAHNIAIASLRPRTFKDLHFLLELQLGHRIRQLGERTFEGLGQ 339
 QY 292 LRELWLNGTHLSTLPAAAFRNLSGLQTLGLTRNPRLSALPRGVFQGLRERLRLV----- 344
 Db 340 LEVLTLNDNQITEVRVGAFLGNVAVMNLSGN-CLRSLSPERVFQGLDKLHSLHLEHSCL 398
 QY 345 ---ALHT-----NALAELRDDALRGLQVSLRHNRRLRAIPRTLFRNLS 387
 Db 399 GHVRLHTEAGLSGLRLFLRDNSISSIEEQSLAGLSELLDLDTTNRLLTHLPRQLFQGLG 458
 QY 388 SL 389
 Db 459 HL 460

Search completed: November 29, 2004, 13:12:03
 Job time : 29 secs

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 13:05:08 ; Search time 25 seconds
 (without alignments)
 1497.133 Million cell updates/sec

Title: US-09-775-803-12_COPY_1_389
 Perfect score: 1938
 Sequence: 1 MIRSALLSAVLLRAQQPFP.....LRHNRLRALPRTLFRNLSSL 389
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_79:*

1: Pirl:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	76.2	560	2 A60164	platelet membrane insulin-like growth factor-binding protein - human
2	419.5	21.6	605	2 A41915	insulin-like growth factor-binding protein - human
3	416.5	21.5	536	2 A34901	insulin-like growth factor-binding protein - human
4	413.5	21.3	605	2 JC5239	insulin-like growth factor-binding protein - human
5	382.5	19.7	603	2 JC1282	insulin-like growth factor-binding protein - human
6	382	19.7	603	2 JC6128	insulin-like growth factor-binding protein - human
7	379.5	19.6	1531	2 T42218	insulin-like growth factor-binding protein - human
8	367.5	19.0	907	2 JG0193	insulin-like growth factor-binding protein - human
9	366	18.9	1523	2 T13953	insulin-like growth factor-binding protein - human
10	345	17.8	1389	2 T13852	insulin-like growth factor-binding protein - human
11	339	17.5	907	2 JE0176	insulin-like growth factor-binding protein - human
12	337	17.4	1051	2 T13174	insulin-like growth factor-binding protein - human
13	337	17.4	1385	2 T13887	insulin-like growth factor-binding protein - human
14	327	16.9	1091	2 A58532	insulin-like growth factor-binding protein - human
15	319.5	16.5	1469	2 A36665	insulin-like growth factor-binding protein - human
16	319.5	16.5	1480	2 A36665	insulin-like growth factor-binding protein - human
17	317.5	16.4	707	2 JC7763	insulin-like growth factor-binding protein - human
18	317	16.4	361	2 A53860	insulin-like growth factor-binding protein - human
19	313	16.2	1066	2 T15864	insulin-like growth factor-binding protein - human
20	311.5	16.1	622	2 JC7973	insulin-like growth factor-binding protein - human
21	310	16.0	575	2 T29972	insulin-like growth factor-binding protein - human
22	306.5	15.8	1134	1 A29944	insulin-like growth factor-binding protein - human
23	305.5	15.8	1097	2 A29943	insulin-like growth factor-binding protein - human
24	305	15.7	961	2 T23395	insulin-like growth factor-binding protein - human
25	302	15.6	610	2 T23836	insulin-like growth factor-binding protein - human
26	301	15.5	594	2 T23841	insulin-like growth factor-binding protein - human
27	300	15.5	789	2 T28714	insulin-like growth factor-binding protein - human
28	300	15.5	1355	2 T28715	insulin-like growth factor-binding protein - human
29	298.5	15.4	1119	2 AD1822	insulin-like growth factor-binding protein - human

garp precursor - h	3.0	296	15.3	662	2	S42799
densin-180 - rat	3.1	295.5	15.2	1495	2	T31434
hypothetical protein	3.2	291	15.0	680	2	T19339
hypothetical protein	3.3	290.5	15.0	738	2	T19338
disease resistance	3.4	288.5	14.9	1112	2	T10504
hypothetical protein	3.5	286.5	14.8	572	2	T30947
hypothetical protein	3.6	286.5	14.8	613	2	A88684
hypothetical protein	3.7	286	14.8	526	2	C84552
leucine-rich alpha	3.8	283.5	14.6	312	1	NBHUA2
Ras-binding protein	3.9	283.5	14.6	559	2	T42998
cell-surface molec	4.0	269	13.9	682	2	A49121
connectin precursor	4.1	266	13.9	682	2	A43318
gene flightless-I	4.2	265.5	13.7	1256	2	S60461
disease resistance	4.3	264.5	13.6	1016	2	T30553
hypothetical protein	4.4	253	13.1	1232	2	T05322

ALIGNMENTS

garp precursor - h	3.0	296	15.3	662	2	S42799
densin-180 - rat	3.1	295.5	15.2	1495	2	T31434
hypothetical protein	3.2	291	15.0	680	2	T19339
hypothetical protein	3.3	290.5	15.0	738	2	T19338
disease resistance	3.4	288.5	14.9	1112	2	T10504
hypothetical protein	3.5	286.5	14.8	572	2	T30947
hypothetical protein	3.6	286.5	14.8	613	2	A88684
hypothetical protein	3.7	286	14.8	526	2	C84552
leucine-rich alpha	3.8	283.5	14.6	312	1	NBHUA2
Ras-binding protein	3.9	283.5	14.6	559	2	T42998
cell-surface molec	4.0	269	13.9	682	2	A49121
connectin precursor	4.1	266	13.9	682	2	A43318
gene flightless-I	4.2	265.5	13.7	1256	2	S60461
disease resistance	4.3	264.5	13.6	1016	2	T30553
hypothetical protein	4.4	253	13.1	1232	2	T05322

garp precursor - h	3.0	296	15.3	662	2	S42799
densin-180 - rat	3.1	295.5	15.2	1495	2	T31434
hypothetical protein	3.2	291	15.0	680	2	T19339
hypothetical protein	3.3	290.5	15.0	738	2	T19338
disease resistance	3.4	288.5	14.9	1112	2	T10504
hypothetical protein	3.5	286.5	14.8	572	2	T30947
hypothetical protein	3.6	286.5	14.8	613	2	A88684
hypothetical protein	3.7	286	14.8	526	2	C84552
leucine-rich alpha	3.8	283.5	14.6	312	1	NBHUA2
Ras-binding protein	3.9	283.5	14.6	559	2	T42998
cell-surface molec	4.0	269	13.9	682	2	A49121
connectin precursor	4.1	266	13.9	682	2	A43318
gene flightless-I	4.2	265.5	13.7	1256	2	S60461
disease resistance	4.3	264.5	13.6	1016	2	T30553
hypothetical protein	4.4	253	13.1	1232	2	T05322

garp precursor - h	3.0	296	15.3	662	2	S42799
densin-180 - rat	3.1	295.5	15.2	1495	2	T31434
hypothetical protein	3.2	291	15.0	680	2	T19339
hypothetical protein	3.3	290.5	15.0	738	2	T19338
disease resistance	3.4	288.5	14.9	1112	2	T10504
hypothetical protein	3.5	286.5	14.8	572	2	T30947
hypothetical protein	3.6	286.5	14.8	613	2	A88684
hypothetical protein	3.7	286	14.8	526	2	C84552
leucine-rich alpha	3.8	283.5	14.6	312	1	NBHUA2
Ras-binding protein	3.9	283.5	14.6	559	2	T42998
cell-surface molec	4.0	269	13.9	682	2	A49121
connectin precursor	4.1	266	13.9	682	2	

Query Match 21.6%; Score 419.5; DB 2; Length 605;
 Best Local Similarity 33.0%; Pred. No. 2.1e-21;
 Matches 126; Conservative 58; Mismatches 155; Indels 43; Gaps 6;

Qy 38 GSVAH--IAEGLPINTHILLFMRMDQGILRNHSESGMTVILQRLMLSDSHISAIIDPGTF 94
 Db 140 GTFAHTPALLASLGLSNN---RLSRLEDGL---FEGLGLWDLNLGWNNSLAVLPDAAF 190

Qy 95 NDLVKLTKLRLTRNKISRLPRAILDKMWVLEQILFELDHNAALRDLDQNLFQQLRNQELGLN 154
 Db 191 RGLGSLRELVLAGNRLAYLQPALFSGLAELREILDSRNALRAIKANVFVQLPRLQRLYLD 250

Qy 155 QNQLSFLPANLFLSSLRNLTHLPK---GILLGAQVKLEKLLLYSNQLTSV 210
 Db 251 RNLIAAVAPGAFPLGLKALRWLQLSHNRVAGLLEDTPFGLLGLRV---LRLSHNAIASLR 306

Qy 211 SGLLSNLGAALTTELRLERNHLRSVAPGAFDRLGNLSSLTLSGNLLESLPPALFLHVSSV 270
 Db 307 PRTFKDLHFLLEELQLGLQLEVLTLDHNQLQEVRAFLGLTNAV 366

Qy 271 LTLFENPLEELPDVLFEGEMAGLRELWLNGTHLSTLPAAAFRNLSGLQTGLLTRN---- 324
 Db 367 MNLSCGNCLRNLPQVERGLGKHLHLEGSCLGRIPHTFTGLSGLRRFLKDNGLGLVIE 426

Qy 325 -----PRLSALPRGVFQGLRELRVIALHTNALAELRDDALRGLGHLRQ 367
 Db 427 EQSLWGLAELLELDLTSNQLTHLPFLRPLFQGLGKLEYLLSRNRLAELPADAALGPLQRAFW 486

Qy 368 VSLRHNLRALPRTLERNLSSL 389
 Db 487 LDVSHNRLEAALPNSLLAPLGRL 508

RESULT 3
 A34901 lysine carboxypeptidase (EC 3.4.17.3) 83K chain - human
 C;Species: Homo sapiens (man)
 C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004
 C;Accession: A34901
 R;Tan, F.; Weerasinghe, D.K.; Skidgel, R.A.; Tamei, H.; Kaul, R.K.; Roninsson, I.
 J.Biol. Chem. 265, 13-19, 1990
 A;Title: The deduced protein sequence of the human carboxypeptidase N high molecular weight chain
 A;Reference number: A34901; MUID:90094386; PMID:2378615
 A;Accession: A34901
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-536 <TAN>
 A;Cross-references: UNIPROT:P22792; GB:J05158; NID:9179935; PIDN:AAA51921.1; PIDN:AAA51921.2
 C;Genetics:
 A;Gene: GDB:ACBP
 A;Cross-references: GDB:127893
 A;Map position: 6q25.3-6q26
 C;Keywords: hydrolase; metallo-carboxypeptidase
 F;1/77-100/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F;1/101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F;1/125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F;1/149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F;1/173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F;1/197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F;2/221-244/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F;2/245-268/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F;2/269-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F;2/293-316/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F;3/317-340/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F; 341-364/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>

Query Match 21.5%; Score 416.5; DB 2; Length 536;

Best Local Similarity 32.3%; Pred. No. 2.9e-21;

Matches 115; Conservative 53; Mismatches 163; Indels 25; Gaps 3;

Db 22 PTKCKVVRDAAQCGGGSVAAHI-----AELGLPTNLTHILLF 58

Db 2 PMGCDCFVQEVF-CSDEELATVPLDIPPYTKNIIFVETSFTTLETRAFGSNPNLTKVFL 60

Qy 59 RMDQGILRNHSFSGMTVLQRLMLSDSHISAI DPGTFNDLVKLTLLTRNKISRLPRAIL 118

Db 61 DTQLCQFRPDAFGGLPRLE DLEVTGSSFLNSTNIFSNLTSGLKLTLMFALPEGLF 120

Qy 119 DKMVLLEQQLFDHNA LRDQNLFOQLRNLQELGLNQNQQLSFLPANL FSSLRELKLLDLS 178

Db 121 QHLAALESLHQLQGNQQLQALP RRLFQPLTHLKTLLNAQNLLAQQLPEELFHPLTSLQTLKLS 180

Qy 179 RNNNLTHLPKG LIGAQVKLEKLLYSNQLTSDSGLLSNLGALT ERLERNHLSRVA PAGAF 238

Db 181 .NNALSGLPLQGVFGKGSIQLQELFLDSNNNISELPQVFSQQLFCLLERLWLQRNAITHLPLSIF 240

Qy 239 DRLGNLSSLTLSGNNI JESLPALFLHVSSVSRLTLEENP LEELPDVLFEGMAGLREILWN 298

Db 241 ASLGNLTFLSLQWMLRVLPAGLPAHTPCLVGLSLTHNQLET VTEGTFAHLSNLRSLMLS 300

Qy 299 GTHLSTLPAAAFRNLSGLOTLGLTRNPRLSALPRGVFOGLRERVLALHTNALAEL 354

Db 301 YNAITHLPA GIFR DLEELV KLYLGSN-NLTALHPALFQNL SKLELLSLSKQNQ LTTL 355

RESULT 5

Db 336 GLGQLEVLTLDHNQLOEVKVGAFGLGTNVAVMNLPEQVFRGLGK LHS LHEG 395

Qy 300 THLSTLPAAAFRNLSGLOTLGLTRN-----PRLSALPRGVFO 336

Db 396 SCLGRIRPHTFEAGL SGLRRLFLKDNGLQSLWGLAELLELDLTSNQ LTHLPHQLFQ 455

Qy 337 GLRELRVLA LHTNALAELR DDLRGLGHLRQVSLRHNRLRALPRTLFRNLSSL 389

Db 456 GLGKLEYL LLSHNR AELPDA DGLPQRAFVLDVSHNRL EALPGSLLA S LGRU 508

Qy RESULT 5

Db JC1282 insulin-like growth factor-binding protein acid labile chain precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: JC1282

R;Dai, J.; Baxter, R.C.

Biochem. Biophys. Res. Commun. 188, 304-309, 1992

A;Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac

A;Reference number: JC1282; MUID: 93038676; PMID: 1384485

A;Accession: JC1282

A;Molecule type: mRNA

A;Residues: 1-603 <DAI>

A;Cross-references: UNIPROT:P35859; GB:S46785; NID:9258002; PID: AAB23770.2; PID: g570593

A;Experimental source: liver

A;Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205

F;1-27/Domain: signal sequence #status predicted <SIG>

F;28-603/Domain: insulin-like growth factor binding protein, acid labile chain #status F;

F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 19.7%; Score 382.5; DB 2; Length 603;

Best Local Similarity 28.7%; Pred. No. 7.5e-19;

Matches 121; Conservative 64; Mismatches 182; Indels 55; Gaps 7;

Db 21 CPKTKCKCVVRD-----AAQCGGGSVAAHI E LGLPTNLTHILLFRMDQGILTRNHSFSGMTV 75

Db 41 CPVACTCSHDDYTDELSVFCCSKNLTHLP-DIPVSTRALWLDGNNLSSIPSAAFQNLSS 99

Qy 76 LQRMLSDSHISAI DPGTFNDLVKLTLLTRNKISRLPRAILDKMV LLE QLFDHNA LR 135

Db 100 LDFLNLQGSWRLSLEPOA LIGLQNLYYLHLERNRLLRNLAVGLFTHTPSLASLSSNLLG 159

Qy 136 DLDQNLFQQLRNQ LQELGLNQNQLSFLPANL FSSLRELKLLDLSRNRLTHLPKG LLLGAQVK 195

Db 160 RLEEGL FQGLSHI WDNLGNWSL VVLPDVTFOGLGNLH E LVAGNKLTYLQPAL FCGLGE 219

Qy 196 LEKULLYSNQ L TSVDG SLLNGLGALTE RLERNHLSRVA PAGAF-----DRLG 242

Db 220 LRELDLSRNALRSVKA N FVHLPLRQLYLDRLNLITAVAPGAFLGKMKALRWLDLSHNRVA 279

Qy 243 NLSS-----LTL SGNLLES LPPAFLHVSSVSRLTLE RLKLLDLSRNRLTHLPKG LLLGAQVK 291

Db 280 GLMEDTFPGLLGLLAVRLA HNAIASL RPTF KDLHFL EQLGHNR IRQ LGERTFEGL GQ 339

Qy 292 IRELWLNGTHLSTL PAAAFRNLSGLOTLGLTRNPRLSALPRGVFOGLRERVLV-----

Db 340 LEVLTNDNQITEVRVGAFSGL FENAVMNLSGN-CLRS LPERVFOGLDKLHS LHLHSCL 398

Qy 345 --ALHT-----N A L A E L R D D A L R G L G H L R Q V S L R H N R L R A L P T L F R N L S 387

Db 399 GHVRLHTEAGL SGLRRLF RDNSI S SIEEQSLAGLSELLELDLTTNR LTHLPQLFQGLG 458

Qy 388 SL 389

Db 459 HL 460

RESULT 6

Db JC6128 insulin-like growth factor binding complex acid labile chain - mouse

C;Species: Mus musculus (house mouse)

C;Date: 23-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

Qy 240 RLGNLSSLTLSGNNLLES LPPAFLHVSSVSRLTLE RLKLLDLSRNRLTHLPKG LLLGAQVK 299

C;Accession: JC6128
 R;Boisclair, Y.R.; Seto, D.; Hsieh, S.; Hurst, K.R.; Ooi, G.T.
 Proc. Natl. Acad. Sci. U.S.A. 93, 10028-10033, 1996
 A;Title: Organization and chromosomal localization of the gene encoding the mouse acid 1
 A;Reference number: JUC6128; MUID:96413591; PMID:8816745
 A;Accession: JC6128
 A;Molecule type: DNA
 A;Residues: 1-603 <BOI>
 A;Cross-references: UNIPROT:P70389; GB:U66900; PIDN:AAB17270.1; PID:gi16216
 C;Comment: This protein is a serum protein and it is of the ternary complex in the physi
 C;Genetics:
 A;Gene: als
 A;Map position: 17

Query Match 19.7%; Score 382; DB 2; Length 603;
 Best Local Similarity 32.3%; Pred. No. 8.1e-19;
 Matches 120; conservative 55; Mismatches 162; Indels 34; Gaps 5;

Db 43 IABELGLPTNLTHILLFRMDQGILRNHSFSGMTVQLRMLSDSHISAIIDPGTENDLVLKLT 102
 Qy 148 LASLSQLGN 148
 Db 103 LRULTRNKISRLFRAILDKMVVLLEQLFLDHNALRDQNLFOQLRNQELGLQNQLSFLP 162
 Qy 199 LVLAGNKLTYLQPAALLCGLGELELDLSRNALRSVKANVFIHLPRLQKLYLDRLNLITAVA 258
 Qy 163 ANLFSSLRELKLLDLSRNITLHLPK---GLLGAQVK-----LEK 198
 Db 259 PRAFLGMKALRWLDLSHNRAVAGLLEDTPGFLGLHVVLRAHNAITSLRPRTFKDOLHFLEE 318
 Qy 199 LILYNSNQLTSVDSGLSNLGAETELRLERNHLRSVAPGAFDRLGNLSSLTISGNLLESLP 258
 Db 319 LQLGHNRIRQLGEKTFEGLQGLEVLTLNDNQIHEVKVGAFFGLENVAVMNLSGNCNCLRSLP 378
 Qy 259 PAFLFHVSSVSRLTFLFENPLEELPDVLFEMAGLRELWLNTHSTLVLKLTLEL 318
 Db 379 EHVPQGLGRHLHSCLGRIRLHTFAGLSGLRFLRDNSISSIEQSLAGLSELLE 438
 Qy 319 LGLTRNPRLSALPRGVFOGLRELRLVLAHNTNAELRDDALRGGLHRLRQVSRLRNRLRAL 378
 Db 439 LDLTAN-OLTHLPRQLFQGLGQLEYLLSNNNQLTWMSEDVLGPLQRAFWLDLSHNRLETP 497
 Qy 379 PRTLFPRNLSSL 389
 Db 498 AEGLFSSLGRL 508

RESULT 8
 JG0193
 G protein-coupled receptor FEX - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 23-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
 C;Accession: JG0193
 R;Bermeij, G.; Methner, A.; Schallier, H.C.; Hermans-Borgmeyer, I.
 Biochem. Biophys. Res. Commun. 254, 273-279, 1999
 A;Title: Identification of a novel seven-transmembrane receptor with homology to glycopr
 A;Reference number: JG0193; MUID:99121227; PMID:9920770
 A;Accession: JG0193
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-907 <HER>

Query Match 19.0%; Score 367.5; DB 2; Length 907;
 Best Local Similarity 27.4%; Pred. No. 1.4e-17;
 Matches 117; Conservative 79; Mismatches 154; Indels 77; Gaps 11;

Db 6 LLSAVLALLR-----AQPPCPKTKCVRDAAQCSGGSVAHI--AELG---LP 49
 Db 8 MLLSLLALLQLVAAGSSPGPDAIPRGCPSPHCHC-----ELDGRMLLRRVDCSDLGLSELP 61
 Qy 50 TNLTHILLFRMDQGILRNHSFSGMTVQLRMLSDSHISAIIDPGTENDLVLKLTLEL 109
 Db 62 SNLSVFTSY-----LDLSMNNISQIPASLLHRLCFILEELRLAGNA 101
 Qy 110 ISRLPRAILDKMVVLLEQLFLDHNALRDQNLFOQLRNQELGLQNQLSFLPANLFSSL 169
 Db 102 LTHIPKGAFTGLHSLKVLMLQNNQLRKVPEEALQNLRSLSQSLRDNHISYVPPSCFSGL 161
 Qy 170 RELKLLDLSRNITLHLPKGLGAQVKLELLYSNQLTSVDSGLLSNLGALTELRERNH 229
 Db 230 LRSVAPGAFDRLGNLSSLTLSGNLLESLPPAL-----FLHVS 266
 Db 222 IHSLGKKCEDGLHSLETLDNNNIDEFTPAITKTLNSNLKELGFHSNNNRSIPERAFCVGNP 281
 Qy 267 SVSLRLTPENPLEELPDVLFEMAGLRELWLNGLQTLPLAFAFRNLSG--LQLQLLT 322
 Db 282 SLITIHFYDNPPIQFVGVSAFOHLPELRLTLNGASHITEFP----HLTGTATLESLTLT 336
 Qy 323 RNPRLSALPRGVFOGLRELRLVLAHNTNAELRDDALRGGLHRLQVSRLRNRLRALPRTL 382
 Qy 21 CPKTKCKVVRDAAQCSGGSVAHIELGLPTNLTHILLFRMDQGILRNHSFSGMTVQLRML 80

Db 337 -GAKISSLPOAVCDOLPNLQQVLDLSYNLLEDL--PSLGGCQKLQKIDLRHNEIYEIKGST 393
 QY 383 FRNLSSL 389
 Db 394 FQQLFNL 400

RESULT 9
 T13953 MEGF5 protein - rat
 N;Alternate names: slit protein homolog
 C;Species: Rattus norvegicus (Norway rat)
 C;Accession: T13953 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A;Reference number: Z14126; MUID:98360089; PMID:9693030
 A;Accession: T13953
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: mRNA
 A;Residues: 1-1389 <ELD>
 A;Cross-references: UNIPROT:Q24591; EMBL:L23171; NID:9415682; PID:g1019104; PID:AAA7920
 C;Genetics:
 A;Gene: wheeler
 A;Cross-references: FlyBase:FBgn0004364

Query Match 17.8%; Score 345; DB 2; Length 1389;
 Best Local Similarity 30.4%; Pred. No. 8.6e-16;
 Matches 113; Conservative 57; Mismatches 150; Indels 52; Gaps 8;

QY 65 LRNHSFSGMTVLQRLIMLSDSHISAIIDPGTFFNDLVKLRTLRLTRNKI-----110
 Db 136 LHGOSFQGLKESELHGDNNIRQLPEGVWCMSMPSLQLLNLTQNIRSASEFLQFSEKLCA 195

Query Match 18.9%; Score 366; DB 2; Length 1523;
 Best Local Similarity 20.2%; Pred. No. 3.4e-17;
 Matches 120; Conservative 85; Mismatches 180; Indels 210; Gaps 8;

QY 3 RSALLSAVALLRAQP-FPCPKTCKVYRDAAQCSGGSSVAHIAELGLPTNLTHILLFRMD 61
 Db 15 RLLALLALASILSGPPAAACPTKCTCSAA-SVDCHGLGRAVPR-GIPRNAERLDDDRNN 72

QY 62 QGILRNHSFSGMTVLQRLMLSDSHISAIIDPGTFFNDLVKLTKLRLTRNKISRLPRAILDKM 121
 Db 73 ITTRITMDFTGLKRNLRVHLEDNQVSVIERGAFQDLKQLERLRLRNKLNKLQVLPELLFOST 132

QY 122 VILLEQLFDHNALRDLDDQNLFOQLRLNQELGLNNONQSLFLPANLFLSSLRELKLLDLSRNN 181
 Db 133 PKLTRDLSENQIQQIPKRFQVTGVKNLQLDNNNHSICEDGAFRALRDLIELTLMNN 192

QY 182 LTHLPKGLLGAQVKLEKLLLYSNQL---206
 Db 193 ISRILVTSFNHMPKIRTLRHSNHLYCDCHLAWLSDWLRQRRTIGQFTLCMAPVHLRGFS 252

QY 207 -----TSVDSGLLNSLG-----223
 Db 253 VADVQKKEYVCPPHSEAPACNANSLSCPSPACSCSNNIVDCRGKGLTEIPANLPEGIVEI 312

QY 224 RLERNNHRSVAPGAFDRGLNQLSSLTLSLGNLLESLPALFLHVSSVSLRLTFENPLEBLPD 283
 Db 313 RLEQNSIKSIPAGAFIQYKKLRIDISKNQISDIAPDAFQGLKSLTSVLVGNKITEIPK 372

QY 284 VLFEGEMAGLIRELWLNGTHLSTLPAAAFRNLSGL-----QTL 319
 Db 373 GLFDGVLVSLQLLNANKINCLRVNTFQDLQNLNLSSLYDNKLQQTISKGLFAPLQSIQTL 432

QY 320 GLTRNP-----325
 Db 433 HLAQNPFWCDCHLKWLDYLDQDNPPIETSGARCSSPRLANKRISQIKSKKFRCSGSEDYR 492

QY 326 -----RLSALP-----RGV 334
 Db 493 NRSSESECMDLVCPEKCRCEGTIVDCSNQKLSRIPSHLPEYTTDLRLNDNDIAVLEATGI 552

QY 335 FOGLRELRLVLAHTNALABELRDDALRGLGHLRQVSLRHNRLRALPRTLFRNLISSL 389
 Db 553 FKKLDNLRKINSSNNRIKEVREGAFDGAAGVQELMLTGNOLETMHGRMFRGLSGL 607

RESULT 11
 JE0176 orphan G protein-coupled receptor precursor - human
 C;Species: Homo sapiens (man)
 C;Accession: JE0176 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
 R;McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
 Biochem. Biophys. Res. Commun. 247, 266-270, 1998
 A;Title: Identification and cloning of an orphan G protein-coupled receptor of the glycc
 A;Reference number: JE0176; MUID:98308104; PMID:9642114
 A;Accession: JE0176
 A;Molecule type: mRNA
 A;Residues: 1-907 <MCD>
 A;Cross-references: UNIPROT:075473; GB:AF062006; NID:93366801; PID: AAC28019.1; PID:9336
 C;Comment: This protein is a receptor for a novel class of glycoprotein ligands.
 C;Genetics:
 A;Gene: HG38
 A;Map position: 12q22-23

F;1-21/Domain: signal sequence #status predicted <SIG>
 F;562-583/Domain: transmembrane #status predicted <TM1>
 F;594-616/Domain: transmembrane #status predicted <TM2>
 F;639-660/Domain: transmembrane #status predicted <TM3>

F;681-701/Domain: transmembrane #status predicted <TM4>
 F;725-744/Domain: transmembrane #status predicted <TM5>
 F;768-791/Domain: transmembrane #status predicted <TM6>
 F;803-824/Domain: transmembrane #status predicted <TM7>

Query Match 1.3% Score 339; DB 2; Length 907;
 Best Local Similarity 27.7%; Pred. No. 1.3e-15;
 Matches 113; Conservative 69; Mismatches 158; Indels 68; Gaps 10;

Qy 13 LLRAQPFPCKTCKCVVRDAAQCGGGSVAAH--AEIUG--LPTNLTHILLFRMDQGILRN 67
 Db 30 LLRG---CPTHCHC---EPDGRMLLRVDCSDULGSELPSNLSVFTSY---- 70

Qy 68 HSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLKTLRLTRNKISRLPRAILDKRNVILLEQL 127
 Db 71 -----LDSMMNTISQULLPNFLPSLRFLEEURLAGNALTYPKGAFTGLYSLKVL 119

Qy 128 FLDHNAALRDLDDQNLFQQLRNLQELGLQNQNLSPFLPANLFSSLRFLKELDLSRNNNLTHLPLK 187
 Db 120 MLQNNQQLRHVPTEALQNRLSULQSLRNLDAHISYVPPSCFSGLHSIRHLWIDDDNALTEIPV 179

Qy 188 GLIGAQVKLEKLLYSNQLTSDVGSSLNLGALTBLRERNHLSVAPGAFDRLGNLSSL 247
 Db 190 QAERSLSALQAMTALANKIHHIPDYAFGNLSSLVVLHNNRRIHSLGKKCFDGLHSLSTL 239

Qy 248 TLSGNLLESLPPAL-----FLHVSSVSRLTLEFNPBLELLPDV 284
 Db 240 DLNYYNNLDEFTTAIRTLNSNLKELGFSNNNIRSIPERAKFVGNGPSLITIHFYDNPIQFVGRS 299

Qy 285 LFGEMAGLREIWLNGTHLSTLPAAAFRNLSG--LQTIGLGLTRNPRLSALPRGVFQGLREL 341
 Db 300 AFQHLPPELRTLTNGASQT---EFPDLTGATNLSESLTLT-GAQISSLPQTVCNQLPPLN 354

Qy 342 RVLALHTNALAELRDDALRGLGHLRQVSLRHNRLRALPRTLFRNLISSL 389
 Db 355 QVLDLSYNLLEDLPSFSV--CQXLOQKIDRHNTEIYEIKVUDTFQQQLSL 400

RESULT 13
 T13887
 Qy 190 QAERSLSALQAMTALANKIHHIPDYAFGNLSSLVVLHNNRRIHSLGKKCFDGLHSLSTL 239
 C;Species: Drosophila melanogaster
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T13887
 R;Chiang, C.; Beachy, P.A.
 Mech. Dev. 47, 225-239, 1994.
 A;Title: Expression of a novel Toll-like gene spans the parasegment boundary and contributes to the dorsal-ventral axis in Drosophila melanogaster
 A;Reference: PMID:95151581; MUID:217805; PMID:7848870
 A;Accession: T13887
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1385 <CHI>
 A;Cross-references: UNIPROT:Q24591; EMBL:S76155; NID:9913247; PID:9913248; PIDN:AAB33383
 C;Genetics:
 A;Cross-references: FlyBase:FBgn0004364
 A;Note: tlr

Query Match 17.4%; Score 337; DB 2; Length 1385;
 Best Local Similarity 29.8%; Pred. No. 3.1e-15;
 Matches 111; Conservative 57; Mismatches 152; Indels 52; Gaps 8;

Qy 65 LRNHSFSGMTVLQRLMLSDSHISAIDPGTFNDLVKLKTLRLTRNKI----- 110
 Db 136 LHGSFQGLKELSEJHGDDNNIRQLPEGVNCMPSLQIUNLTONRIRSAEFLGFSEKLCIA 195

Qy 111 -SRLPRA-----ILD----- 142
 Db 196 GSALSNANGAVSGGSELTQTLDSFNELRSLPDAWGASRLRLQTLSSLOHNNNISTLAEPNAL 255

Qy 143 QQLRNQBLQGLNQNQLSFLPANLFSLREKLIDLSRNNNLTHLPKGGLGAQVKLEKLL-- 200
 Db 256 AGLSSLRVLNNTSYNHLVSLPSEAFAGNKEELRELHQLQNDLYELPKGLLH--RLEQLLVL 312

Qy 201 -LYSNQLTS--VDSGLLQNLQNSFLPANLFSLREKLIDLSRNNNLTHLPKGGLGAQVKLEKLL-- 257
 Db 313 DLSGNQLTDHHVDNSTEAGLRLIVLNLSNNNATRIGSKTFKELYFLQILDMRNNSIGHI 372

Qy 258 PPAFLHVSSVSRLTLEFNPBLELLPDVLEGEMAGLREIWLNGTHLSTLPAAAFRNLSSLQ 317
 Db 373 EGAFLPLYNLHTNLAENRLHTLDDNRIFNGLYVVLTKLTLNMMNLVSIVESQAFRNCSDLK 432

Qy 318 TLGLTRNPRLSALPRGVFOGLRELEVLALHTNALAELRDDALRGLGHLRQVSLRHNRLRA 377
 Db 433 ELDSLSS-N-QLTVPEAA-QDLISMLKTLQDLDGENQISEFKNNNTFRNLNQLTGLRLIDRIGN 490

Qy 378 LPTLFRNLSSL 389
 Db 491 ITVGMFQDLPLRL 502

Qy 163 ANLFSSIREIUKLDDLSRNNNLTHLPKGGLGAQV----- 202
 Db 406 EKAFEKVTLLEELDLSYNSLTELPRDIFNGTTSLISILHKYNTPNGDILHFGTKDQLQQLDLS 465

Qy 203 SNQTSVDSQGLLSNLGALTBLRERNHLSVAPGAFDRLGNLSSLTLSGNLLESLPPALF 262
 Db 466 FNSIVQVHHSMDKMPGLTNLNKGNGGIKKIQPDSTLXNLRHIDLSINDLQISGMLF 525

Qy 263 LHSSVSRSLTLPENP-LLELPDVLFEGEMAG--LREIWLNGTHLSTLPAAAFRNLSSLQ 318
 Db 526 FKNSELDVIRLNDNPRLSOLPDTGFLSYSGEFTVYLDISNCIAIGPLGHKAFTSTMPEHLTT 585

Qy 319 LGLTRNPRLSALPRGVFQGLREI RLQVSLRHNRLBAL 378
 Db 586 LKLAWN-NINHLPREI FTGLHKLJIDDDLSNNLITRMDNLTEFLMDNGELTKLSSLAGNPISRL 644

Qy 379 PRTLFRNLSSL 389
 Db 645 SVRLFLPLHQL 655

RESULT 14

A58532 glial cell membrane glycoprotein LIG-1 precursor - mouse
C;Species: *Mus musculus* (house mouse)
C;Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 09-Jul-2004
R;Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.
J. Biol. Chem. 271, 22522-22527, 1996
A;Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in A;Accession number: A58532; MUID: 96394313; PMID: 8798419
A;Accession: A58532
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-1091 <SUZ>
A;Cross-references: UNIPROT:P70193; GB:D78572; NID:91545806; PIDN:BA111416.1; PID:915458
F;36-61/Domain: Proteoglycan amino-terminal homology <PAH>
F;71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F;334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F;358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F;385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F;409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F;440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>
Query Match Score 327; DB 2; Length 1091;
Best Local Similarity 28.4%; Pred. No. 1.1e-14;
Matches 114; Conservative 78; Mismatches 182; Indels 28; Gaps 11;

Qy 6 LLSAVALL-----RAQP-FPCPKTCKCVRDAAQCSGGSVAHIAABLGLPT----N 51
Db 19 LLWLLLLLQWPESAGAQARPAPCAAAC-T-AGNSLDCCSGRGLATLPR-DLPSWTRSLN 76

Qy 52 LTHILLFRMDQGLLRNHSFGMTVLQRMLLSDSHISAIIDPGTFNDLVKLKTLTRNPKIS 111
Db 77 LSYNRSEIDSA-----AFEDLTNLQQEYLNNSNELTAI-PSLGTASIGVVSLFLQHNKIL 130

Qy 112 RLPRAILDKMVLLEQFLDHNALRDLDQNLFQQLRNQNQLSFLPANLFFSSL-R 170
Db 131 SVDGSQLKSYLSIEVLDLSSNNITEIRSSCFPNGLRIRELNLA NRISILESGAFDGLSR 190

Qy 171 EIKLUDLSRNNLTHLPKGLLGAQVKLQLLQNQLTSVDSGILLSNLGALTELRLERNHL 230
Db 191 SLLTLLSKRNITOLPVKAFLKP-RLTQDLDNRNRLIRLIEGLTFOGLDSLEVLQRNNI 249

Qy 231 RSVAPGAFDRLGNLSSLTSGNLLSLLPAPLFLHVSSEVSRLTFENPLEELPDVLFGEMA 290
Db 250 SRLTDGAFWGLSKMHLLEYNSLVEVNSGSLYGLTAHQHLHISNNISIRQDGWSSFCQ 309

Qy 291 GLREWLWNGTHLSTLPAAAFRNLSSLGQTLGLTRNPRLSALPRGVFQGLRELRVLA 350
Db 310 KLLHELLSFLNNLTTRLDEEESTAELSSLSLIRLSSH-AISHIAEGAFKGLKSLRVLDDHNE 368

Qy 351 LAELRDD--ALRGLGLRQVSRLRNLALPRTLFRNLSS 389
Db 369 ISGTRIEDSGAFTGLDNLSSKLTLEFGNKKSVAKRAFGSLESL 410

Qy 287 GEMAGLRELWLWNGTHLSTLPAAAFRNLSSLGQTLGLTRNPRLSALPRGVFQGLRELRVLA 346
Db 343 SSFRRRLRIDLSNNNTSRIAHDLSGLKQLTTLVLYGN-KIKDLPGVFKGLGSLRLLL 401

RESULT 15

B36665 slit protein 2 precursor - fruit fly (Drosophila melanogaster)
C;Species: *Drosophila melanogaster*
C;Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
C;Accession: B36665

